University of Alberta

Cyclo-Stationary Signal Analysis & Its Applications in System Identification

by

Jiandong Wang

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**

Department of Electrical and Computer Engineering

Edmonton, Alberta
Spring 2007

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-29764-3
Our file Notre référence
ISBN: 978-0-494-29764-3

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

University of Alberta

Library Release Form

Name of Author: Jiandong Wang

Title of Thesis: Cyclo-Stationary Signal Analysis & Its Applications in System Identification

Degree: Doctor of Philosophy

Year this Degree Granted: 2007

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Jiandong Wang

Date: Jan. 19, 2007

Abstract

System identification deals with the problem of building mathematical models of dynamical systems based on the observed data. Most contemporary studies in this field have a fundamental assumption: the observed data are stationary, which means that statistical characteristics of the data do not change with time. The thesis is motivated by an "ambitious" thought: is it possible to remove or weaken this assumption so that the knowledge in the field can be advanced? The answer is positive by introducing cyclo-stationary signals, which exhibit periodicity in their mean, correlation, and spectral descriptions.

The thesis consists of two parts. The first part studies cyclo-stationary signal analysis, including cyclo-period estimation, cyclo-statistic estimation and cyclo-spectral theory; they provide the second part with powerful computational tools and build up a solid theoretical background. The second part is to exploit cyclo-stationarity in system identification, including finite-impulse-response modeling for errors-in-variables/closed-loop systems, and blind identification of Hammerstein nonlinear systems. The main contributions achieved are briefly described as follows:

1. Cyclo-period estimation: A new method, named as the variability method, is proposed to estimate the cyclo-period of a discrete-time cyclo-stationary signal. Properties of the variability method are analyzed and compared with three existing cyclo-period estimation methods via simulation and real-life examples.

2. Cyclo-statistic estimation: We summarize the existing estimators of the time-varying mean/correlation and cyclic correlation/spectrum, and supplement a new cyclic spectrum estimator: the blocking-based estimator, and discuss implementation issues of these estimators.

3. Cyclo-spectral theory: Two problems in the spectral theory of discrete-time cyclo-stationary signals are studied: (i) four types of the cyclospectrum representation are presented and their interrelationships are explored; (ii) the problem of the cyclospec-

trum transformation is attacked in the framework of multirate systems using the blocking technique as a systematic solution.

4. Finite-impulse-response modeling for errors-in-variables/closed-loop systems: A complete study of the cyclic correlation analysis, which consistently estimates finite-impulse-response models, is developed including the time- and frequency-domain statistical performance of the models.

5. Blind identification of Hammerstein nonlinear systems: A new blind approach is proposed for identification of Hammerstein nonlinear systems by exploiting input's piecewise constant property. In a real-time laboratory experiment, the proposed approach is successfully applied to modeling of a magneto-rheological damper.

To my beloved wife, Qingkun, who has always been the power source leading me "over the top".

Acknowledgements

I would like to take the opportunity to thank all the people and organizations who have helped me in completing the thesis.

First of all, a hearty gratitude is owed to my supervisors, Dr. Tongwen Chen and Dr. Biao Huang, for their outstanding guidance and support all the way. Their efforts have been the most important factors in the creation of the thesis. Second, I am truly grateful to Dr. Akira Sano for his great help and support, which lead to the success of our research collaboration and the wonderful experience for me being a visiting researcher at the Keio University, Japan.

I am obliged to the members in my committee for their input and interest: Dr. Er-Wei Bai, Dr. Fraser Forbes and Dr. Qing Zhao, and the other professors in the Systems Group and the Computer Process Control Group for their assistance and help in my entire graduate study at the UofA: Dr. Alan Lynch, Dr. Horacio J. Marquez, Dr. Edward S. Meadows and Dr. Sirish L. Shah. I would also like to thank our research collaborators at the Matrikon Inc., Canada: Dr. David Shook and Dr. Jianping Gao.

I am thankful to my labmates including postdoctoral fellows both in the Department of Electrical and Computer Engineering and in the Department of Chemical and Materials Engineering for the friendship, company and entertainment they provided in these years.

The research contained in the thesis was supported by the Natural Sciences and Engineering Research Council of Canada, the Alberta Ingenuity Fund, the Informatics Circle of Research Excellence, the Izaak Walton Killam Trusts, and the Japan Society for the Promotion of Science — many thanks to them!

Finally, I would like to thank my parents, parents-in-law, sister, sister-in-law and brothers-in-law from the bottom of my heart for their constant support and encouragement for years.

Contents

1 Introduction 1

2 Preliminary 4
- 2.1 Stationary Signals . . . 4
- 2.2 Cyclo-Stationary Signals . . . 4
- 2.3 Blocked Representation . . . 5

3 Cyclo-Period Estimation 7
- 3.1 Introduction . . . 7
- 3.2 Variability Analysis . . . 9
  - 3.2.1 Relationship between Two Estimators . . . 9
  - 3.2.2 Largest Variability at $n = lp$ . . . 13
  - 3.2.3 Generalization Based on the Time-Varying Mean . . . 15
- 3.3 Properties of the Variability Method . . . 15
  - 3.3.1 Algorithm and Examples . . . 16
  - 3.3.2 Properties of the Variability Method . . . 23
- 3.4 Conclusion . . . 25
- 3.5 Appendix . . . 25

4 Cyclo-Statistic Estimation 30
- 4.1 Introduction . . . 30
- 4.2 Cyclo-Mean Estimators . . . 31
- 4.3 Cyclo-Correlation Estimators . . . 34
- 4.4 Cyclic Spectrum Estimators . . . 36
- 4.5 Cyclo-Stationarity and Quasi-Stationarity . . . 40
- 4.6 Conclusion . . . 42

5 Cyclo-Spectral Theory 43
- 5.1 Introduction . . . 43
- 5.2 Cyclospectrum . . . 46
- 5.3 Blocking Operator . . . 50
- 5.4 Cyclo-Stationary Signals in Multirate Systems . . . 53
  - 5.4.1 Cyclo-Stationarity of the Output . . . 53
  - 5.4.2 Cyclospectrum of the Output . . . 55
- 5.5 Conclusion . . . 59

6 FIR Modeling for EIV/Closed-Loop Systems **60**
6.1 Introduction 60
6.2 Joint Cyclo-Stationary Signals 63
6.3 Cyclic Correlation Analysis 64
6.3.1 Definition 64
6.3.2 Statistical Performance 67
6.4 Discussion 71
6.4.1 Induced Cyclo-Stationarity 71
6.4.2 Existing Cyclo-Stationarity 73
6.5 Numerical Illustration 74
6.6 Conclusion 79

7 Blind Identification of Hammerstein Systems **81**
7.1 Introduction 81
7.2 Problem Description 83
7.3 Identification of Linear Dynamics 85
7.3.1 Two Groups of Basic Equations 85
7.3.2 Estimation of $A(q)$ 87
7.3.3 Estimation of $B(q)$ 89
7.3.4 Determination of n_a, n_b and τ 90
7.4 Inner Signal Estimation 91
7.5 Theoretical Analysis 93
7.6 Algorithm and Simulation 95
7.7 MR Damper Modeling 99
7.7.1 Hammerstein Model 99
7.7.2 Experiment Setup 100
7.7.3 Experiment Result 101
7.8 Conclusion 105

8 Conclusion **107**

Bibliography **109**

List of F s

1.1 A schematic diagram of the thesis structure 3

3.1 $\sigma_v^2 = 0$, 20 independent trials (a) The variability method ($n = 16l$), (b) Hurd-Gerr's method ($\lfloor 1024/64 \rceil = 16$), (c) Martin-Kedem's method ($\lfloor 1024/64 \rceil = 16$), (d) Dandawaté-Giannakis's method ($\lfloor 1024/64 \rceil = 16$) with $P_F = 1\%$ (dashline). 17

3.2 $\sigma_v^2 = 0$, sample mean of 100 trials (a) The variability method, (b) Hurd-Gerr's method, (c) Martin-Kedem's method, (d) Dandawaté-Giannakis's method. 18

3.3 $\sigma_v^2 = 5$ (a) The variability method ($n = 16l$), (b) Hurd-Gerr's method, (c) Martin-Kedem's method, (d) Dandawaté-Giannakis's method ($\lfloor 4096/256 \rceil = 16$). 18

3.4 Ill-cyclo-stationarity in correlation at lag 0 (a) The variability method based on $\hat{R}_{y_n^{(k)} y_n^{(k)}}(0)$, (b) The variability method based on $\hat{R}_{y_n^{(k)} y_n^{(k+1)}}(0)$ ($n = 7l$), (c) Martin-Kedem's method, (d) Hurd-Gerr's method ($\lfloor 1024/146 \rceil = 7$). . . 19

3.5 Cyclo-period inconsistency (a) The variability method based on $m_x(t)$ ($n = 10l$), (b) The variability method based on $R_{xx}(t;0)$ ($n = 5l$), (c) Martin-Kedem's method ($\lfloor 1024/102 \rceil = 10$), (d) Dandawaté-Giannakis's method ($\lfloor 1024/204 \rceil = 5$). 20

3.6 Two cyclo-stationarities (a) The variability method ($n = 12l$), (b) Hurd-Gerr's method ($\lfloor 1024/256 \rceil = 4$, $\lfloor 1024/170 \rceil = 6$), (c) Martin-Kedem's method ($\lfloor 1024/256 \rceil = 4$, $\lfloor 1024/170 \rceil = 6$), (d) Dandawaté-Giannakis's method ($\lfloor 1024/256 \rceil = 4$, $\lfloor 1024/170 \rceil = 6$). 21

3.7 Global irradiance (a) The variability method ($n = 24l$), (b) Hurd-Gerr's method ($\lfloor 4096/170 \rceil = 24$), (c) Martin-Kedem's method ($\lfloor 4096/170 \rceil = 24$), (d) Dandawaté-Giannakis's method ($\lfloor 4096/170 \rceil = 24$). 22

4.1 Estimate the time-varying correlation ($N = 4001$): $r_{xx}(t;\tau)$ (solid), $\hat{r}_{xx}^{(i)}(t;\tau)$ (dotts), $\hat{r}_{xx}^{(ii)}(t;\tau)$ (circles) and the 95% confidence interval of $\hat{r}_{xx}^{(i)}(t;\tau)$ (dash). 35

4.2 Estimate the cyclic spectrum ($N = 4096$): $S_{yy}(k;e^{j\omega})$ (smooth & solid), $\hat{S}_{yy}^{(i)}(k;e^{j\omega})$ (thick solid) and $\hat{S}_{yy}^{(iii)}(k;e^{j\omega})$ (dash). 40

5.1 A linear SISO multirate system 45

5.2 A cascade of upsampler ($\uparrow m$), LTI system (H) and downsampler ($\downarrow n$) . . . 45

5.3 Interrelationships among the four cyclospectra 49

5.4 Blocking a linear SISO multirate system 50

5.5 A blocked SISO multirate system 53

5.6 An equivalent blocked SISO multirate system: $pr = qn$ 55

5.7 A maximally decimated filter bank 58

5.8 Polyphase representation of a filter bank 58

6.1 A general framework including open-loop, EIV, and closed-loop systems . . 61
6.2 FIR modeling of an EIV system from the CCRA ($N = 2000$): the true impulse response coefficients (solid) and the estimates from the CCRA (circle) with the 3-standard deviation band from (6.16) (dash) and that from 100 Monte Carlo simulations (plus). 75
6.3 FIR modeling of an EIV system from the CRA ($N = 2000$): the true impulse response coefficients (solid) and the estimates from the CRA (dot) with two 3-standard deviation band from (6.16) (dash) and that from 100 Monte Carlo simulations (plus). 75
6.4 NRMSEs of the CCRA (circle) and CRA (dot) for different N 76
6.5 NRMSEs of the CCRA (circle) and CRA (dot) for different noise levels . . 76
6.6 FIR modeling of a closed-loop system ($N = 2000$): the true impulse response coefficients (solid), the estimates from the CRA (dot), from the CCRA (circle) with 3-standard deviation band (short-dash), from the JCRA (star) with 3-standard deviation band (long-dash). 79

7.1 A discrete-time Hammerstein system with sampling period h 84
7.2 A sampled-data Hammerstein system 85
7.3 An equivalent slow-rate SIMO FIR model 92
7.4 A graph of the controller output $u(n)$ v.s. the estimated inner signal $\hat{X}(n)$. 98
7.5 A graph of the controller output $u(n)$ v.s. the true inner signal $X(n)$ 98
7.6 Experimental devices 100
7.7 A diagram of the experimental setup 101
7.8 Some enlarged parts of experimental data 102
7.9 Selected experimental data 102
7.10 The estimated inner signal $X(n)$ 103
7.11 The nonlinearity (dots) revealed by $U(n)$ and $\hat{X}(n)$, and the estimated nonlinearity (smooth line) 103
7.12 The measured (solid) and simulated (dotted) damping forces using the estimation data: fitness = 70.5102% 105
7.13 The measured (solid) and simulated (dotted) damping forces using the validation data: fitness = 63.1926% 105

List of Tables

3.1 Sample means of $\hat{R}_{x_n^{(k)}x_n^{(k)}}(0)$ in the 100 trials 12

4.1 Estimate the time-varying mean 33

6.1 NRMSEs of $\hat{\theta}_1$ (CCRA) and $\hat{\theta}_{1,p-1}$ (aggregated CCRA) for $s(t)$ with frequency-band $[0, 1]$ 78
6.2 NRMSEs of $\hat{\theta}_1$ (CCRA) and $\hat{\theta}_{1,p-1}$ (aggregated CCRA) for $s(t)$ with frequency-band $[0.2, 0.8]$ 78

7.1 AIC for different combinations of $\hat{n}_a$ and $\hat{\tau}$ 96
7.2 The variance of $\Delta_w^{(l)}$ 97
7.3 AICs for time delay estimation 97
7.4 Estimated parameters and their standard deviations 97
7.5 Averaged fitnesses and their standard deviations: the standard deviations of F^0_{LSM} and F^0_{INV} are omitted 99
7.6 The AICs under different combinations of $\hat{n}_a$ and $\hat{\tau}$ 103

List of Abbreviations

AIC	Akaike information criterion
AR	autoregressive
ARX	autoregressive with exogenous variables
CCRA	cyclic correlation analysis
CRA	correlation analysis
$(CS)_p$	cyclo-stationary with cyclo-period p
CSPA	cyclic spectral analysis
DCS	degree of cyclo-stationarity
DFS	discrete Fourier series
DFT	discrete Fourier transformation
DTFT	discrete-time Fourier transformation
EIV	errors-in-variables
FFT	fast Fourier transformation
FIR	finite impulse response
IVM	instrumental variable method
JCRA	joint correlation analysis
LPTV	linear periodically time-varying
LSM	least-squares method
LTI	linear time invariant
MPC	model predictive control
MR	magneto-rheological
NRMSE	normalized root-mean-square-error
RBS	random-binary sequence
SIMO	single-input and multiple-output
SISO	single-input and single-output
SPA	spectral analysis
SNR	signal-to-noise ratio
TFR	time frequency representation
WN(0, 1)	white noise with zero mean and unit variance
ZOH	zero-order hold

List of Symbols

$E\{\cdot\}$	expectation
$\bar{E}\{\cdot\}$	generalized expectation: $\bar{E}\{\cdot\} := \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} E\{\cdot\}$
$\mathbb{Z}$	set of integers
$\mathbb{Z}_+$	set of nonnegative integers
L_n	n-fold discrete blocking operator
L_n^{-1}	inverse of L_n
$x(t)$	discrete-time signal
$\underline{x}_n$	n-fold blocked representation of $x(t)$: $\underline{x}_n = L_n x$
$\underline{x}_n^{(k)}$	k-th component of $\underline{x}_n$
$\lfloor \cdot \rceil$	the nearest integer function
$\lfloor \cdot \rfloor$	floor function (the greatest integer function)
$'$	transpose
*	complex conjugate transpose
$\mathrm{Var}(\cdot)$	variance of the operand
$\mathrm{Cov}(\cdot)$	covariance of the operand
q	one sample advance operator: $qx(t) = x(t+1)$
$\gcd(m,n)$	the greatest common divisor of two integers m and n

Chapter 1

Introduction

System identification deals with the problem of building mathematical models of dynamical systems based on the observed data [122, 85]. As dynamical systems are abundant in our environment, the techniques of system identification have a wide application area including engineering and science. Most contemporary studies in this field have a fundamental assumption, namely, the observed data are wide-sense stationary, which basically means that statistical characteristics of the data do not change with time. The thesis is essentially motivated by an "ambitious" thought: is it possible to remove or weaken this assumption so that the knowledge in the field can be advanced? The answer is positive by introducing cyclo-stationary signals.

Discrete-time signals are said to be wide-sense cyclo-stationary, if their correlations and/or means are periodically time-varying sequences [62, 56, 61]. Cyclo-stationary signals often arise from the time-varying nature of physical phenomena such as the weather [20, 91], and more importantly from certain man-made operations, e.g., amplitude modulation, time index modulation, fractional sampling and multirate system filtering [56, 61]. Under these circumstances, exploring cyclo-stationarity is more reasonable and promising than ignoring and treating cyclo-stationary signals as if they were stationary. In particular, system identification often has a freedom of designing identification tests/experiments to make the observed data sufficiently informative. By this freedom, cyclo-stationarity can be readily introduced into system identification by man-made operations when it does not exist naturally.

The fore-mentioned "ambitious" thought is challenging but feasible. As a matter of fact, we have done/seen some contributions towards this direction. In [133, 151], cyclo-stationarity was introduced by sampling outputs faster and updating control inputs slower by zero-order hold, and identifiability was achieved in principle for closed-loop systems

without external excitation. This result overrides a well-known theorem in the system identification theory and brings a new possibility into the horizon. In the field of communication, cyclo-stationarity spurs a substantial breakthrough in the blind identification [142, 83], a problem closely related to system identification. The fast-sampling technique converts stationary communication outputs into cyclo-stationary sequences, whose second-order statistics contain enough information, particularly the phase information, to identify the possibly non-minimum phase communication channels.

The idea of exploiting cyclo-stationarity for system identification[1] possibly originated from [53], where time-difference-of-arrival, namely, an errors-in-variables (EIV) system with only time delay, was attacked by an algorithm that is a prototype of the cyclic correlation analysis (CCRA) studied later in Chapter 6. The CCRA, a technique of estimating finite-impulse-response (FIR) coefficients of linear-time invariant (LTI) systems, was formally presented in [60, 61] without a detailed analysis. The frequency-domain counterpart of the CCRA, namely, the cyclic spectral analysis (CSPA), was proposed in [53, 55] to give asymptotically unbiased frequency-response estimates for EIV systems. The CSPA was generalized for identification of closed-loop systems in [60] and was completed in [7] in the sense of developing the statistical performance (means and variances) of the estimated frequency responses. In [47], frequency responses were estimated from spectral cross-moments and cumulants of high-order cyclo-stationary signals by an algorithm whose computational complexity is comparable to the CSPA. In terms of parametric identification of LTI systems, cyclo-stationarity has received little attention. In [132, 145], the CSPA was taken as the first step to estimate frequency responses, which were treated as data to give parametric models in the second step. By sampling outputs faster and updating control inputs slower, cyclo-stationarity was brought up to achieve identifiability of the direct approach for closed-loop systems without external excitation [133, 151]. Cyclo-stationarity has also occasionally been introduced for nonlinear systems such as Volterra kernels [58, 92], polyperiodic systems [57, 81, 93] and block-oriented nonlinear systems [105, 131, 14]. To summarize, the study of exploiting cyclo-stationarity in system identification is only at the early stage, except that the CSPA has been fully developed [53, 55, 60, 7].

Before exploiting cyclo-stationarity in system identification, we need a solid theoretical background of cyclo-stationary signals and some efficient computational tools. Hence, the

[1]Cyclo-stationarity has received considerable attention for the blind identification during the last two decades; however, the blind identification is very different from system identification in the sense that the former is solely based on outputs. Hence, literatures on the blind identification and other related areas are omitted here — see [116] for a recent comprehensive bibliography on cyclo-stationarity.

first part of the thesis studies the following three topics of cyclo-stationary signal analysis:

1. Cyclo-period estimation — Determine from observed data the cyclo-period of a discrete-time cyclo-stationary signal.

2. Cyclo-statistic estimation — Estimate the first- and second-order statistics of cyclo-stationary signals from observed data.

3. Cyclo-spectral theory — Study the spectral descriptions of cyclo-stationary signals, and the cyclospectrum transformation by linear systems.

The second part of the thesis aims at two specific system identification topics:

4. FIR modeling for EIV/closed-loop systems — Develop a complete study of the CCRA including the statistical performance of the estimated FIR coefficients for EIV/closed-loop LTI systems.

5. Blind identification of Hammerstein systems — Propose a new approach to blind identification of Hammerstein systems, where a static nonlinearity precedes a linear dynamical subsystem.

The overall structure of the thesis is schematically illustrated in Figure 1.1.

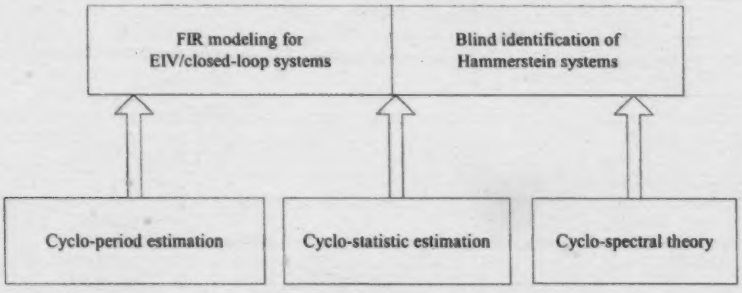

Figure 1.1: A schematic diagram of the thesis structure

The rest of the thesis is organized as follows. Chapter 2 introduces some preliminary concepts such as stationarity and cyclo-stationarity. Each of the above five topics is studied in Chapters 3-7, respectively. Chapter 8 summarizes the thesis by listing its main contributions. The thesis is written in the paper format since most of the chapters are actually published in journals or conferences [152]-[160].

Chapter 2

Preliminary

This chapter reviews some basic concepts of stationary and cyclo-stationary signals with their blocked representation.

2.1 Stationary Signals

A discrete-time signal x is said to be stationary or wide-sense-stationary [100] if its mean is constant,

$$E\{x(t+\tau)\} = E\{x(t)\} =: m_x,\ \forall \tau \in \mathbb{Z}, \tag{2.1}$$

and its autocorrelation depends only on the time difference,

$$E\{x(t+\tau)x^*(t)\} = R_{xx}(\tau),\ \forall \tau \in \mathbb{Z}. \tag{2.2}$$

The power spectrum of x is defined as the discrete-time Fourier transform (DTFT) of the autocorrelation,

$$S_{xx}\left(\mathrm{e}^{j\omega}\right) = \sum_{\tau=-\infty}^{\infty} R_{xx}(\tau)\,\mathrm{e}^{-j\omega\tau}. \tag{2.3}$$

It is well known that when a stationary signal x with power spectrum $S_{xx}\left(\mathrm{e}^{j\omega}\right)$ is passed through an LTI system with transfer function $\hat{G}(z)$, the output y is also a stationary signal with power spectrum [100, 85]

$$S_{yy}\left(\mathrm{e}^{j\omega}\right) = \hat{G}\left(\mathrm{e}^{j\omega}\right) S_{xx}\left(\mathrm{e}^{j\omega}\right) \hat{G}^*\left(\mathrm{e}^{j\omega}\right). \tag{2.4}$$

2.2 Cyclo-Stationary Signals

A discrete-time signal x is said to be cyclo-stationary or strictly cyclo-wide-sense-stationary, if its mean and/or correlation are periodic sequences [62, 56]. In particular, x is called first-order cyclo-stationary [61] if its time-varying mean $m_x(t) := E\{x(t)\}$ is periodic,

$$m_x(t + lp_1) = m_x(t),\ \ \forall t, l \in \mathbb{Z}. \tag{2.5}$$

Similarly, x is second-order cyclo-stationary [61] if its time-varying correlation

$$R_{xx}(t;\tau) := E\{x(t+\tau)x^*(t)\} \tag{2.6}$$

is periodic in t for a fixed τ,

$$R_{xx}(t+lp_2;\tau) = R_{xx}(t;\tau), \forall t, l \in \mathbb{Z}. \tag{2.7}$$

Here p_1 and p_2 are the smallest positive integers such that (2.5) and (2.7) hold, respectively. If $p_1 = p_2 = 1$, (2.5) and (2.7) imply that the mean is time invariant, and the correlation depends on the time difference only; thus, x is wide-sense stationary. In other words, stationary signals can be regarded as cyclo-stationary signals with period 1.

A cyclo-period inconsistency problem occurs frequently; in this case, the periods of $m_x(t)$ and $R_{xx}(t;\tau)$ are different. First, the second-order cyclo-stationarity can arise alone, i.e., $p_1 = 1$ and $p_2 \neq 1$. Second, if the first- and second-order cyclo-stationarities coexist, the two periods may not be the same, i.e., $p_1 \neq p_2$, $p_1 \neq 1$ and $p_2 \neq 1$. Therefore, the cyclo-period p is defined as the least common multiple of p_1 and p_2; x is said to be cyclo-stationary with period p, abbreviated as $(\mathrm{CS})_p$.

Example 2.1 Let us see the cyclo-period inconsistency problem via an example: $x(t) = \cos(2\pi t/4)\, w(t)$, where w is stationary with constant mean m_w and delay-dependent correlation $R_{ww}(\tau)$. The time-varying mean and the time-varying correlation of x respectively are

$$m_x(t) = \cos\left(\frac{2\pi t}{4}\right) E\{w(t)\} = \cos\left(\frac{2\pi t}{4}\right) m_w,$$

and

$$\begin{aligned} R_{xx}(t;\tau) &= \cos\left(\frac{2\pi t}{4}\right)\cos\left(\frac{2\pi(t+\tau)}{4}\right) E\{w(t+\tau)\, w^*(t)\} \\ &= \frac{1}{2}\left[\cos\left(\pi t + \frac{2\pi\tau}{4}\right) + \cos\left(\frac{2\pi\tau}{4}\right)\right] R_{ww}(\tau). \end{aligned}$$

Two cases exist: (i) If $m_w = 0$, then $m_x(t) = 0$ and x is second-order cyclo-stationary only; the cyclo-period is 2. (ii) If $m_w \neq 0$, the first- and second-order cyclo-stationarities coexist, but with different periods 4 and 2, respectively; the cyclo-period is 4. □

2.3 Blocked Representation

Let $x(t)$ be a discrete-time signal defined on $\mathbb{Z}_+$. The n-fold discrete blocking operator L_n is defined as the mapping from a scalar sequence x to a n-dimensional vector sequence $\underline{x}_n$,

where underlining denotes blocking [96, 147][1]:

$$x \mapsto \underline{x}_n = \begin{bmatrix} x_n^{(0)} \\ x_n^{(1)} \\ \vdots \\ x_n^{(n-1)} \end{bmatrix} := \left\{ \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(n-1) \end{bmatrix}, \begin{bmatrix} x(n) \\ x(n+1) \\ \vdots \\ x(2n-1) \end{bmatrix}, \begin{bmatrix} x(2n) \\ x(2n+1) \\ \vdots \\ x(3n-1) \end{bmatrix}, \cdots \right\}. \quad (2.8)$$

In the subsequent chapters, $\underline{x}_n$ may be denoted as $\underline{x}$ for simple notation if the subscript is obvious from the context. The dimension and the sampling period of $\underline{x}_n$ equal n times those of x, but no information is lost in the blocking operation [28]. Thus, $\underline{x}_n$ can be regarded as the blocked representation of $x(t)$. The blocked signal $\underline{x}_n$ has other names in the literature, e.g., the time series representation [56] and the decimated component [61]. The inverse of the blocking operator, L_n^{-1}, is defined as the reverse operation of (2.8); thus, $L_n^{-1}L_n = I$ and $L_n L_n^{-1} = I$, where I denotes the identity system. Important relationships between blocked signals and original signals are: $x(t)$ is $(\mathrm{CS})_p$ if and only if $\underline{x}_p$ is stationary; $x(t)$ and $y(t)$ are jointly $(\mathrm{CS})_p$ if and only if $\underline{x}_p$ and $\underline{y}_p$ are jointly stationary [113, 112, 2, 152].

[1]The blocking in signal processing is also known as lifting in control [78, 28].

Chapter 3

Cyclo-Period Estimation

This chapter[1] presents a new method, named as the variability method, to estimate the cyclo-period of a discrete-time cyclo-stationary signal. The method is essentially based on the time-varying correlation and/or the time-varying mean, whose estimators are associated with some statistics of blocked signals; a plot of variability of these statistics as a function of the blocking operator index visually reveals a periodic pattern, from which the cyclo-period is obtained. Properties of the variability method are analyzed and compared with three existing cyclo-period estimation methods via simulation and real-life examples.

3.1 Introduction

The cyclo-period, defined as the least common multiple of periods of mean and correlation sequences (Section 2.2), is the most fundamental parameter of a cyclo-stationary signal; hence, estimation of the cyclo-period should be regarded as the first step whenever the cyclo-period is required to be known *a priori*. The purpose of this chapter is to present a new method to estimate the cyclo-period from a time series[2], i.e., given a realization of a cyclo-stationary signal x with unknown cyclo-period p,

$$\{x(t)\}_{t=0}^{N-1} := \{x(0), x(1), \cdots, x(N-1)\}, \tag{3.1}$$

how to estimate p?

In the literature, there have been some methods aiming to or being applicable to estimating the cyclo-period. Herbst [70] tested the periodic fluctuation in the variance function of a cyclo-stationary signal via periodogram, which can be adapted to estimate the cyclo-period. Tian [141] inferred the period via a limiting property of sample autocovariances.

[1] The chapter has been published in [153, 156].

[2] The proposed variability method can actually detect cyclo-stationarity because it observes a periodic pattern that is from the periodic variability of the time-varying correlation/mean; even though there are possibilities that time-varying correlation/mean are aperiodic, the possibilities are really dim in practice.

Observing a phenomenon that the period of the mean function may be different from that of the variance function, Martin and Kedem [90] formed a new periodic sequence having the period equal to the least common multiple of periods of the mean and variance functions of the original cyclo-stationary signal, and then detected the period via the periodogram associated with the new sequence. Later, Martin [91] was motivated by some special cyclo-stationary signals having zero-means and unit variances, e.g., a seasonal time series, and formed a stationarity test (not cyclo-period estimation) based on a relationship between the correlation function and the probability of zero-crossing. Hurd and Gerr [72] obtained the cyclo-period from the bispectrum, which was estimated using the two-dimensional periodogram. Dandawaté and Giannakis [37] aimed at the detection of cyclo-stationarity under a broader context, *almost* cyclo-stationary signals, through a statistical χ^2 test based on the cyclic covariance and cyclic spectrum, where the cyclo-period was actually estimated as well. Among these methods, Hurd-Gerr's, Martin-Kedem's, and Dandawaté-Giannakis's methods are more complete than the others and will be compared with our proposed method. Their main ideas and algorithms are summarized in Section 3.5. Besides the above explicitly-related methods, the cyclo-period can also be obtained from those statistic estimators for cyclo-stationary signals that do not assume knowledge of the cyclo-period *a priori*, e.g., the cyclic periodogram in [53, 54, 110, 25]; however, the estimated result may suffer from a phenomenon called the cycle leakage, resulting in a poor resolution; this is due to lack of mechanisms like those in Hurd-Gerr's and Dandawaté-Giannakis's methods, which are specifically designed to eliminate noise effects.

The new method to be proposed, referred to as the variability method, has many attractive features comparing to the other three methods mentioned earlier. First, it is not sensitive to stationary noises while Hurd-Gerr's and Martin-Kedem's methods are. Second, it is equally applicable to different types of ill-cyclo-stationary signals (to be clarified later), while Martin-Kedem's method cannot handle a special ill-cyclo-stationarity. Third, it can deal with the cyclo-period inconsistency problem (see Section 2.2), while Hurd-Gerr's and Dandawaté-Giannakis's methods cannot. Finally, it provides the best resolution and is easier to use than other methods. On the other hand, the price of these attractive features is that after estimating the cyclo-period correctly, the variability method cannot detect the simultaneous existence of two or more cyclo-stationarities, while the other three methods may be capable of the detection. It is also worthy to mention that the variability method works for cyclo-stationary signals exclusively, while Dandawaté-Giannakis's method is applicable to *almost* cyclo-stationary signals where the cyclo-period may not be an integer

but a real-valued number.

The rest of the chapter is organized as follows. Section 3.2 establishes the theoretical foundation of the variability method. Properties of the variability method are illustrated through examples in Section 3.3, followed by a conclusion in Section 3.4. Section 3.5 is the appendix summarizing the main ideas and algorithms of Hurd-Gerr's, Martin-Kedem's, and Dandawaté-Giannakis's methods.

3.2 Variability Analysis

The purpose of this section is to build a theoretical foundation for the variability method. The exposition is based on the time-varying correlation, because the first-order cyclo-stationarity sometimes may not exist (see Section 2.2) and the methodology to be used is equally applicable if the time-varying correlation is replaced by the time-varying mean. First, the time-varying correlation of a cyclo-stationary signal is connected with some statistics of a blocked signal formed from the original signal (Theorem 3.1). Next, those statistics of blocked signals are proved to have the largest variability measured by sample variance when the blocking operator index is an integer multiple of the cyclo-period (Theorem 3.2). Finally, the above two conclusions are generalized to the case that the variability method is based on the time-varying mean (Theorems 3.3 and 3.4).

3.2.1 Relationship between Two Estimators

We present an estimator of $R_{xx}(t;\tau)$ defined in (2.6) and an estimator of "correlation"[3] of $x_n^{(k)}$ defined in (2.8), and explore the relationship between the two estimators. In the sequel, $x(t)$ is assumed to be real-valued and univariate for the sake of an easier presentation.

Given $\{x(t)\}_{t=0}^{N-1}$ in (3.1), an estimator of $R_{xx}(t;\tau)$ is,

$$\hat{R}_{xx}(t;\tau) = \frac{1}{\lfloor N/p \rfloor} \sum_{k=0}^{\lfloor (N-\tau)/p \rfloor - 1} x(kp+t)\, x(kp+t+\tau), \tag{3.2}$$

where $0 \le t \le p-1$ and $0 \le \tau \le N-p$ [3, 94, 61]. For fixed t and τ, $u(k) := x(kp+t)$ and $v(k) := x(kp+t+\tau)$ are the components of the blocked signal $\underline{x}_p$ and are jointly stationary [113, 112]; hence, $\hat{R}_{xx}(t;\tau)$ is the same as the well-known correlation estimator defined for stationary signals and is asymptotically unbiased and consistent, i.e., $\hat{R}_{xx}(t;\tau)$ converges to $R_{xx}(t;\tau)$ as $N \to \infty$.

[3] Correlation is an appropriate term only if n is an integer multiple of p.

The k-th component of the blocked signal $\underline{x}_n = L_n x$ defined in (2.8) is[4]

$$x_n^{(k)} = \{x(k), x(n+k), x(2n+k), \cdots, x((\lfloor N/n \rfloor - 1)n + k)\}. \tag{3.3}$$

Even though $x_n^{(k)}$'s are jointly stationary only at $n = lp$ [113, 112], the following statistics are still computed using the correlation estimator defined for stationary signals,

$$\hat{R}_{x_n^{(k_1)} x_n^{(k_2)}}(d) = \frac{1}{\lfloor N/n \rfloor} \sum_{r=0}^{\lfloor N/n \rfloor - \tau - 1} x(rn + k_1)\, x(rn + k_2 + dn), \tag{3.4}$$

where $0 \le \tau \le \lfloor N/n \rfloor - 1$, $0 \le k_1 \le n-1$, $k_1 \le k_2$, and $x_n^{(k_2)} = q^{\left\lfloor \frac{k_2}{n} \right\rfloor} x_n^{\left(k_2 - n\left\lfloor \frac{k_2}{n} \right\rfloor\right)}$.

Theorem 3.1 *Given a fixed τ and $\{x(t)\}_{t=0}^{N-1}$ in (3.1), the two estimators in (3.2) and (3.4) are connected as*

$$\hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0) \simeq \frac{1}{\bar{p}} \sum_{r=0}^{\bar{p}-1} \hat{R}_{xx}(rn + k; \tau), \tag{3.5}$$

where $\bar{p} = p/\gcd(p, n)$ and $0 \le k \le n-1$. The difference is ignorable for a large N and reduces to zero under some configurations, e.g., $\tau = 0$ and N is a common multiple of n and p.

Proof of Theorem 3.1: Let $c = \gcd(p, n)$, i.e.,

$$p = \bar{p} \cdot c, \quad n = \bar{n} \cdot c, \tag{3.6}$$

which imply that $\bar{p}$ and $\bar{n}$ are coprime or relatively prime. From (3.4),

$$\begin{aligned}
\hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0) &= \frac{1}{\lfloor N/n \rfloor} \sum_{l=0}^{\lfloor N/n \rfloor - \tau - 1} x(ln + k)\, x(ln + k + \tau) \\
&\simeq \frac{1}{\lfloor N/n \rfloor} \sum_{r=0}^{\bar{p}-1} \sum_{m=0}^{\left\lfloor \frac{\lfloor N/n \rfloor - \tau}{\bar{p}} \right\rfloor - 1} x(mp\bar{n} + rn + k)\, x(mp\bar{n} + r\bar{n} + k + \tau) \\
&= \frac{1}{\lfloor N/n \rfloor} \sum_{r=0}^{\bar{p}-1} \left\lfloor \frac{\lfloor N/n \rfloor}{\bar{p}} \right\rfloor \hat{R}_{xx}(rn + k; \tau) \\
&\simeq \frac{1}{\bar{p}} \sum_{r=0}^{\bar{p}-1} \hat{R}_{xx}(rn + k; \tau).
\end{aligned}$$

Here the second approximate equality is achieved by changing l to $(m\bar{p} + r)$ and by replacing $\bar{p}n$ with $p\bar{n}$ (see (3.6)); it holds for a large N comparing to τ, n and p, and becomes an

[4]The data points $x(\lfloor N/n \rfloor n)$, $x(\lfloor N/n \rfloor n + 1)$, $\cdots$, $x(N)$ may be discarded in the blocking operation.

equality for some configurations, e.g., $\tau = 0$ and N is a common multiple of n and p; the third equality is from the definition in (3.2); the last approximate equality holds for a large N and reduces to an equality if $\bar{p} = 1$ or N is a common multiple of n and p. □

Remark: In practice, $\tau = 0$ or 1 is usually enough to estimate the cyclo-period for most cyclo-stationary signals; in addition, $R_{xx}(t;\tau)$ may decay rapidly with respect to τ, which implies that choosing a large τ is not desirable. If the two estimators in Theorem 3.1 are different, they converge to each other in a rate proportional to $1/N$, as implied by the above proof.

Let us look at Theorem 3.1 in detail.

- If $n = lp$ for $l \in \mathbb{Z}_+$, (3.5) reduces to

$$\begin{aligned} \hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0) &\simeq \frac{1}{\bar{p}} \sum_{r=0}^{\bar{p}-1} \hat{R}_{xx}(rlp + k; \tau) \\ &= \frac{1}{\bar{p}} \sum_{r=0}^{\bar{p}-1} \hat{R}_{xx}(k; \tau) \\ &= \hat{R}_{xx}(k; \tau), \end{aligned} \tag{3.7}$$

  where the second equality is from the periodicity in (2.7). Eq. (3.7) says that if $n = lp$, computing the statistics of the blocked signal, $\hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0)$, actually estimates $R_{xx}(k;\tau)$, the time-varying correlation of the original signal.

- Another special case is that p and n are coprime or relatively prime. A result from the number theory is needed (e.g., Theorem 2 on page 44 in [45]):

  Lemma 3.1 *If p and n are relatively prime integers, k is an arbitrary integer, and $\{r_0, r_1, \cdots, r_{p-1}\}$ is a complete residue system modulo p, then*

$$\{nr_0 + k, nr_1 + k, \cdots, nr_{p-1} + k\}$$

  is a complete residue system modulo p.

  Since $\{0, 1, \cdots, p-1\}$ is a complete residue system modulo p, Lemma 3.1 gives that if p and n are coprime, i.e., $\bar{p} = p$, (3.5) becomes

$$\begin{aligned} \hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0) &\simeq \frac{1}{p} \sum_{r=0}^{p-1} \hat{R}_{xx}(rn + k; \tau) \\ &= \frac{1}{p} \sum_{t=0}^{p-1} \hat{R}_{xx}(t; \tau). \end{aligned} \tag{3.8}$$

n	$\hat{R}_{x_n^{(0)}}(0)$	$\hat{R}_{x_n^{(1)}}(0)$	$\hat{R}_{x_n^{(2)}}(0)$	$\hat{R}_{x_n^{(3)}}(0)$	$\hat{R}_{x_n^{(4)}}(0)$	$\hat{R}_{x_n^{(5)}}(0)$	$\hat{R}_{x_n^{(6)}}(0)$	$\hat{R}_{x_n^{(7)}}(0)$
2	1.9916	0.9988						
3	1.4900	1.4937	1.5011					
4	3.9832	1.0007	0	0.9970				
5	1.5020	1.4914	1.4908	1.5035	1.4884			
6	1.9829	1.0001	2.0042	0.9974	1.9875	0.9990		
7	1.4901	1.4909	1.4905	1.5000	1.4955	1.4996	1.4998	
8	3.9862	1.0002	0	0.9936	3.9801	1.0011	0	1.0005

Table 3.1: Sample means of $\hat{R}_{x_n^{(k)}x_n^{(k)}}(0)$ in the 100 trials

Eq. (3.8) says, if p and n are coprime, $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)$ is invariant to k, i.e.,

$$\hat{R}_{x_n^{(0)}x_n^{(\tau)}}(0) = \hat{R}_{x_n^{(1)}x_n^{(1+\tau)}}(0) = \cdots = \hat{R}_{x_n^{(n-1)}x_n^{(n-1+\tau)}}(0),$$

and is equal to the average of the whole set of

$$\left\{\hat{R}_{xx}(k;\tau)\right\}_{k=0}^{p-1} := \left\{\hat{R}_{xx}(0;\tau), \hat{R}_{xx}(1;\tau), \cdots, \hat{R}_{xx}(p-1;\tau)\right\}. \tag{3.9}$$

- Besides the above two cases, (3.5) says that $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)$ is the average of some subsets of $\left\{\hat{R}_{xx}(k;\tau)\right\}_{k=0}^{p-1}$.

Example 3.1 We verify Theorem 3.1 via a numerical example. It is adapted from an example in [72]:

$$x(t) = \left[1 + \cos\left(\frac{2\pi t}{4}\right)\right] w(t),$$

where w is white noise with zero-mean and unit variance, abbreviated as WN$(0,1)$. The time-varying mean is a constant,

$$m_x(t) = \left[1 + \cos\left(\frac{2\pi t}{4}\right)\right] E\{w(t)\} = 0; \tag{3.10}$$

the time-varying correlation is $R_{xx}(t;\tau) = 0$, $\forall \tau \neq 0$ and

$$\begin{aligned} R_{xx}(t;0) &= \left[1 + \cos\left(\frac{2\pi t}{4}\right)\right]^2 E\{w^2(t)\} \\ &= 1 + 2\cos\left(\frac{2\pi t}{4}\right) + \cos^2\left(\frac{2\pi t}{4}\right). \end{aligned} \tag{3.11}$$

Thus, $R_{xx}(0;0) = 4$, $R_{xx}(1;0) = 1$, $R_{xx}(2;0) = 0$, $R_{xx}(3;0) = 1$, and $R_{xx}(t+4;0) = R_{xx}(t;0)$. In other words, x is second-order cyclo-stationary with period 4. A Monte Carlo simulation of 100 independent trials with $N = 10^4$ is implemented to compute $\hat{R}_{x_n^{(k)}x_n^{(k)}}(0)$ for different n. Sample means of $\hat{R}_{x_n^{(k)}x_n^{(k)}}(0)$ in the 100 trials are shown in the Table 3.1[5]; the simulation results support Theorem 3.1. □

[5] $\hat{R}_{x_n^{(k)}}(0)$ is a short notation of $\hat{R}_{x_n^{(k)}x_n^{(k)}}(0)$.

3.2.2 Largest Variability at $n = lp$

For a fixed τ, the sequence

$$\left\{\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)\right\}_{k=0}^{n-1} := \left\{\hat{R}_{x_n^{(0)}x_n^{(\tau)}}(0), \hat{R}_{x_n^{(1)}x_n^{(1+\tau)}}(0), \cdots, \hat{R}_{x_n^{(n-1)}x_n^{(n-1+\tau)}}(0)\right\} \tag{3.12}$$

is to be shown having the largest variability at $n = lp$, $\forall l \in \mathbb{Z}_+$, which is intuitively true:

- If $n = lp$, (3.7) gives $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0) = \hat{R}_{xx}(k;\tau)$; cyclo-stationarity implies that $\hat{R}_{xx}(k;\tau)$'s are generally not same for all integers k, i.e., some variability exists.
- If p and n are coprime, (3.8) says that $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)$'s are the same for all integers k, i.e., the variability is zero.
- Besides the above two cases, $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)$ is the average of some subset of $\left\{\hat{R}_{xx}(k;\tau)\right\}_{k=0}^{p-1}$; hence, there may exist some variability, but the variability is generally smaller than that at $n = lp$ because of the averaging effect.

Before proving this heuristic argument, we need a new definition for some special cyclo-stationary signals.

Definition 3.1 *A discrete-time $(CS)_p$ signal x is called **ill-cyclo-stationary in correlation at lag τ** if $R_{xx}(t;\tau)$'s are the same for all integers t, i.e.,*

$$R_{xx}(0;\tau) = R_{xx}(1;\tau) = \cdots = R_{xx}(p-1;\tau).$$

Ill-cyclo-stationary signals arise frequently in practice. For example, the output of a discrete-time zero-order hold (ZOH) is ill-cyclo-stationary in correlation at lag $\tau = 0$ (see Example 3.3); the output of an upsampler is ill-cyclo-stationary in correlation at lag $\tau \neq lp$ for $l \in \mathbb{Z}$. Ill-cyclo-stationarity conceptually has some overlaps with the first- and second-order cyclo-stationarities. For instance, the signal in Example 3.1 is both second-order cyclo-stationary only and ill-cyclo-stationary in correlation at lag τ, $\forall \tau \neq 0$.

Theorem 3.2 *If x is not ill-cyclo-stationary in correlation at lag τ, the sequence $\left\{\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)\right\}_{k=0}^{n-1}$ defined in (3.12) has the largest variability at $n = lp$, $\forall l \in \mathbb{Z}_+$, in terms of the sample variance*

$$\frac{1}{n}\sum_{k=0}^{n-1}\left(\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0) - \frac{1}{n}\sum_{k=0}^{n-1}\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)\right)^2. \tag{3.13}$$

Proof of Theorem 3.2: First, the sample mean of the sequence $\left\{\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)\right\}_{k=0}^{n-1}$ is invariant to n and is equal to the average of the whole set of $\left\{\hat{R}_{xx}(t;\tau)\right\}_{t=0}^{p-1}$ defined in (3.9), shown as follows:

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1}\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0) &= \frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{\bar{p}}\sum_{r=0}^{\bar{p}-1}\hat{R}_{xx}(rn+k;\tau)\\ &= \frac{1}{p\bar{n}}\sum_{m=0}^{p\bar{n}-1}\hat{R}_{xx}(m;\tau)\\ &= \frac{1}{p}\sum_{t=0}^{p-1}\hat{R}_{xx}(t;\tau),\end{aligned}$$

where the first equality uses (3.5); the second equality is obtained from (3.6) and by changing $(rn+k)$ to a new variable m; the last equality is from the periodicity of $\hat{R}_{xx}(t;\tau)$. Therefore, without loss of generality, the sample mean of $\left\{\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)\right\}_{k=0}^{n-1}$ is assumed to be zero in order to ease the illustration.

Second, if $n = lp$, the sample variance of $\left\{\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)\right\}_{k=0}^{n-1}$ is the same as $\frac{1}{p}\sum_{t=0}^{p-1}\hat{R}_{xx}^2(t;\tau)$, because $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0) = \hat{R}_{xx}(k;\tau)$ (see (3.7)) and $\hat{R}_{xx}(k;\tau)$ is periodic in k with period p. Thus, the difference between the sample variance at $n = lp$ and that at $n \neq lp$ is

$$\begin{aligned}&\frac{1}{p}\sum_{t=0}^{p-1}\hat{R}_{xx}^2(t;\tau) - \frac{1}{n}\sum_{k=0}^{n-1}\hat{R}^2_{x_n^{(k)}x_n^{(k+\tau)}}(0)\\ =\ &\frac{1}{p}\sum_{t=0}^{p-1}\hat{R}_{xx}^2(t;\tau) - \frac{1}{n}\sum_{k=0}^{n-1}\left[\frac{1}{\bar{p}}\sum_{r=0}^{\bar{p}-1}\hat{R}_{xx}(rn+k;\tau)\right]^2\\ =\ &\frac{1}{n\bar{p}^2}\left\{\bar{p}\sum_{t=0}^{\bar{n}p-1}\hat{R}_{xx}^2(t;\tau) - \sum_{k=0}^{n-1}\left[\sum_{r=0}^{\bar{p}-1}\hat{R}_{xx}(rn+k;\tau)\right]^2\right\}\\ =\ &\frac{1}{n\bar{p}^2}\sum_{k=0}^{n-1}\left\{\bar{p}\sum_{r=0}^{\bar{p}-1}\hat{R}_{xx}^2(rn+k;\tau) - \left[\sum_{r=0}^{\bar{p}-1}\hat{R}_{xx}(rn+k;\tau)\right]^2\right\}\\ >\ &0,\end{aligned}$$

where the first equality is from (3.5), the second uses the periodicity of $\hat{R}_{xx}(t;\tau)$ and (3.6), the third is obtained by changing the variable t to $(rn+k)$, and the last inequality is from the non-ill-cyclo-stationarity assumption and Chebyshev Sum Inequality (see e.g., [67]): For a univariate real-valued sequence $\{x_0, x_1, \cdots, x_{n-1}\}$,

$$n\sum_{i=0}^{n-1}x_i^2 \geq \left(\sum_{i=0}^{n-1}x_i\right)^2,$$

where the equality holds iff $x_0 = x_1 = \cdots = x_{n-1}$. □

3.2.3 Generalization Based on the Time-Varying Mean

The generalization is mainly motivated by the cyclo-period inconsistency problem introduced in Section 2.2. The exposition in Sections 3.2.1 and 3.2.2 reveals that the methodology is applicable to the case based on the time-varying mean in a completely parallel fashion; hence, proofs of the theorems are omitted. First, a time-varying mean estimator for $\{x(t)\}_{t=0}^{N-1}$ is,

$$\hat{m}_x(t) = \frac{1}{\lfloor N/p \rfloor} \sum_{k=0}^{\lfloor N/p \rfloor - 1} x(kp+t), \tag{3.14}$$

where $0 \le t \le p-1$. Analogously to (3.4), the "mean" estimator for $x_n^{(k)}$ defined in (3.3) is,

$$\hat{m}_{x_n^{(k)}} = \frac{1}{\lfloor N/n \rfloor} \sum_{r=0}^{\lfloor N/n \rfloor - 1} x(rn+k), \tag{3.15}$$

where $0 \le k \le n-1$. Next, the following are the counterparts of Theorem 3.1, Definition 3.1 and Theorem 3.2.

Theorem 3.3 *Given $\{x(t)\}_{t=0}^{N-1}$ in (3.1), the two estimators in (3.14) and (3.15) are connected as*

$$\hat{m}_{x_n^{(k)}} \simeq \frac{1}{\bar{p}} \sum_{r=0}^{\bar{p}-1} \hat{m}_x(rn+k),$$

where $\bar{p} = p/\gcd(p,n)$ and $0 \le k \le n-1$. The difference is ignorable for a large N and reduces to zero if N is a common multiple of n and p.

Definition 3.2 *A discrete-time $(CS)_p$ signal x is called* ***ill-cyclo-stationary in mean*** *if $m_x(t)$'s are the same for all integers t, i.e.,*

$$m_x(0) = m_x(1) = \cdots = m_x(p-1).$$

Theorem 3.4 *If x is not ill-cyclo-stationary in mean, the sequence*

$$\left\{\hat{m}_{x_n^{(k)}}\right\}_{k=0}^{n-1} := \left\{\hat{m}_{x_n^{(0)}}, \hat{m}_{x_n^{(1)}}, \cdots, \hat{m}_{x_n^{(n-1)}}\right\}$$

has the largest variability at $n = lp$, $\forall l \in \mathbb{Z}_+$, in terms of the sample variance

$$\frac{1}{n}\sum_{k=0}^{n-1}\left(\hat{m}_{x_n^{(k)}} - \frac{1}{n}\sum_{k=0}^{n-1}\hat{m}_{x_n^{(k)}}\right)^2. \tag{3.16}$$

3.3 Properties of the Variability Method

This section first presents simulation and real-life examples in a comparative manner to analyze the variability method and confirm its performance. Second, the properties of the variability method are summarized with the underlying rationales.

3.3.1 Algorithm and Examples

Algorithm: Given $\{x(t)\}_{t=0}^{N-1}$ in (3.1), the cyclo-period p can be estimated by the following steps:

1. Start with $n = 2$, i.e., block x by L_2 as defined in (2.8). Select an integer lag τ (usually 0 or 1) if the variability method is to be based on the time-varying correlation.
2. Compute $\hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0)$ in (3.4) or/and $\hat{m}_{x_n^{(k)}}$ in (3.15) for $k = 1, 2, \cdots, n$, and the sample variance in (3.13) or/and that in (3.16).
3. Repeat Steps 1 and 2 by increasing n until a reasonable number n_{max}, where $n_{max} \geq lp$ for some positive integer l. A good rule of thumb is $50 n_{max} \leq N$ [22][6].
4. Plot the sample variance as a function of the blocking operator index n, where the largest peaks display a periodic pattern at $n = lp$, i.e., p is the smallest integer among the cluster of the largest peaks.

Remark: The cyclo-period p is visually picked in the above algorithm. Sometimes it would be practically sensible to obtain p by an embeddable algorithm, namely, an algorithm detects the periodicity of and estimates the period of the sample variance computed in the above step 2. Standard techniques such as spectral estimation [106], auto-correlation [106] and wavelet transforms [26] may not work well for such a short-length sequence (the length of the sample variance is n_{max}). The periodic-subspace projection/decomposition method [117, 97] is more suitable here.

Let us apply the variability method with the other three existing methods (see Section 3.5) to different examples, starting from investigating the effects of stationary noises.

Example 3.2 The example is from [72] (see Eq. (14) therein):

$$x(t) = \left[1 + \cos\left(\frac{2\pi t}{16}\right)\right] w(t) + v(t),$$

where w is WN$(0,1)$, v is WN$(0, \sigma_v^2)$, and w and v are mutually independent. It is seen from Example 3.1 that x is second-order cyclo-stationary with period $p = 16$. First, to study the noise-free performance, i.e., $\sigma_v^2 = 0$, a Monte Carlo simulation of 100 independent trials is implemented with the parameter configurations as follows:

- *The variability method:* Since only $R_{xx}(t;0)$ is non-zero (see Example 3.1), let $\tau = 0$; the maximum blocking operator index $n_{max} = 40$; the data length N in (3.1) is 1024, the same as the fast Fourier transformation (FFT) lengths in the other three methods.

[6]The ratio 50 has to be higher if the noise level is high.

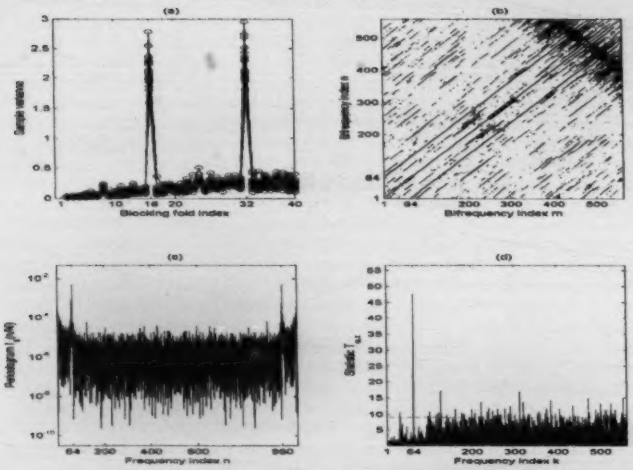

Figure 3.1: $\sigma_v^2 = 0$, 20 independent trials (a) The variability method ($n = 16l$), (b) Hurd-Gerr's method ($\lfloor 1024/64 \rfloor = 16$), (c) Martin-Kedem's method ($\lfloor 1024/64 \rfloor = 16$), (d) Dandawaté-Giannakis's method ($\lfloor 1024/64 \rfloor = 16$) with $P_F = 1\%$ (dashline).

- *Hurd-Gerr's method:* The FFT length T in (3.22) is 1024, the parameter M in (3.23) is 64 and the confidence level P_0 in (3.24) is 95%.

- *Martin-Kedem's method:* The parameter K in (3.25) is 20, the FFT length T in (3.27) is 1024 and the parameter M in (3.26) is chosen as $(T + K)$.

- *Dandawaté-Giannakis's method:* The smoothing window W_L in (3.28) and (3.29) is a Kaiser window with parameter 1 and length $L = 61$, the FFT length T in (3.30) is 1024 and the false alarm level P_F is 1%.

The first 20 samples and the sample mean of the 100 trials are shown in Figures 3.1 and 3.2, respectively. All four methods estimate the cyclo-period $p = 16$ correctly. Second, the effect of a stationary noise is investigated by increasing the noise variance σ_v^2 to 5. To cope with such a high level noise, the data length and all FFT lengths are increased to 4096, while the other parameter are unchanged[7]. Figure 3.3 shows the results of one typical sample of 100 trials. The variability method and Dandawaté-Giannakis's method correctly estimate the cyclo-period, while Hurd-Gerr's and Martin-Kedem's methods fail. The periodic patterns in Figures 3.1-(a) and 3.3-(a) are very similar, which shows that the variability method is insensitive to noise; see also Example 3.6. □

[7]In the subsequent examples, the parameter configurations of the four methods are the same as those used to generate Figures 3.1 and 3.2 unless stated explicitly.

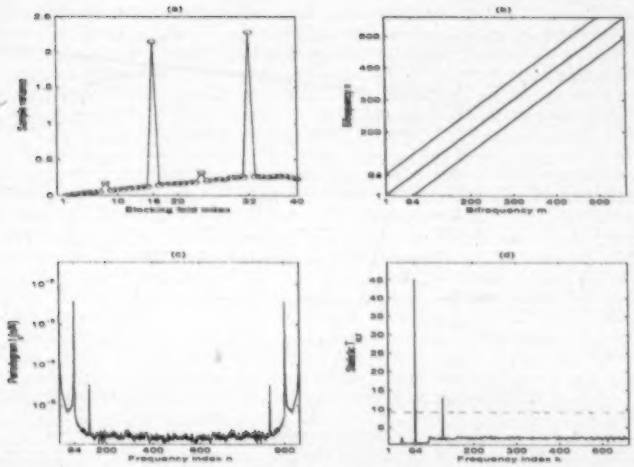

Figure 3.2: $\sigma_v^2 = 0$, sample mean of 100 trials (a) The variability method, (b) Hurd-Gerr's method, (c) Martin-Kedem's method, (d) Dandawaté-Giannakis's method.

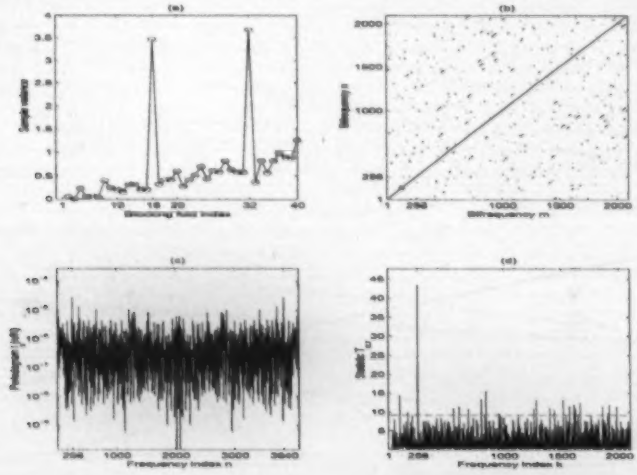

Figure 3.3: $\sigma_v^2 = 5$ (a) The variability method ($n = 16l$), (b) Hurd-Gerr's method, (c) Martin-Kedem's method, (d) Dandawaté-Giannakis's method ($\lfloor 4096/256 \rfloor = 16$).

It seems that for most second-order cyclo-stationary signals, $\hat{R}_{x_n^{(k)} x_n^{(k)}}(0)$ provides enough information for cyclo-period estimation. However, if the ill-cyclo-stationarity in correlation at lag $\tau = 0$ arises (see Definition 3.1), the variability method needs to rely upon $\hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0)$ for $\tau \neq 0$. Let us look at an interesting example. A discrete-time ZOH with an integer factor p, H_p, takes an input sequence x and produces an output sequence

$$y(t) = x\left(\left\lfloor \frac{t}{p} \right\rfloor\right). \tag{3.17}$$

The discrete-time ZOH is an useful operator, e.g., in the scenario of fast sampling the output of a closed-loop system [151]. Eq. (3.17) implies that if y is blocked by L_p,

$$\underline{y}_p = L_p y = \left[\begin{array}{cccc} y_p^{(0)} & y_p^{(1)} & \cdots & y_p^{(p-1)} \end{array} \right]',$$

where $y_p^{(k)}$'s are all identical to x (see (2.8)); hence, $\underline{y}_p$ is stationary iff x is. In addition, the power spectrum of $\underline{y}_p$ is not pseudo-circulant [113], which implies that if x is stationary, y is $(\mathrm{CS})_p$ and $p \neq 1$. Moreover, due to the uniformity of $y_p^{(k)}$ and Theorem 3.1, $R_{yy}(t;0)$'s are the same for all integers t, i.e., y is ill-cyclo-stationary in correlation at lag 0.

Example 3.3 We estimate the cyclo-period of the output of a discrete-time ZOH H_7 driven by a stationary input x, where x is generated by passing a WN(0, 1) signal through an autoregressive (AR) filter with poles $0.45 \pm j0.35$. Figure 3.4-(a) displays the variability of $\hat{R}_{y_n^{(k)} y_n^{(k)}}(0)$ as a function of n, where an aperiodic pattern occurs, as expected. Figure 3.4-(b) is based on $\hat{R}_{y_n^{(k)} y_n^{(k+1)}}(0)$, i.e., $\tau = 1$, where the cyclo-period $p = 7$ is correctly estimated. Since Martin-Kedem's method is based on the mean and variance (corresponding to $\tau = 0$) functions , it can not deal with such an ill-cyclo-stationary signal, which is confirmed by Figure 3.4-(c). Both Hurd-Gerr's and Dandawaté-Giannakis's methods succeed in the estimation, for they are capable of using all the second-order statistical information; Figure 3.4-(d) shows the result obtained by Hurd-Gerr's method. □

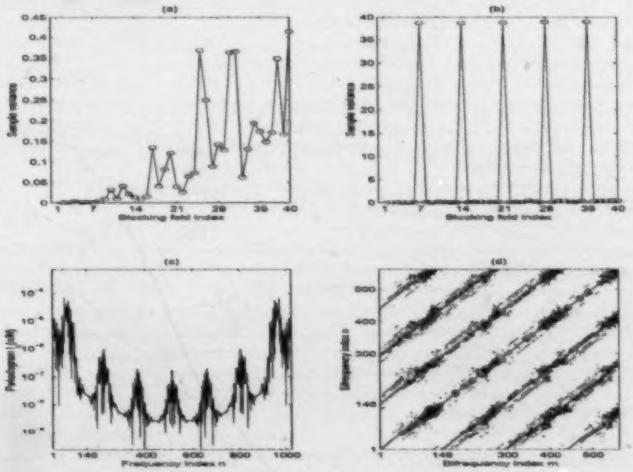

Figure 3.4: Ill-cyclo-stationarity in correlation at lag 0 (a) The variability method based on $\hat{R}_{y_n^{(k)} y_n^{(k)}}(0)$, (b) The variability method based on $\hat{R}_{y_n^{(k)} y_n^{(k+1)}}(0)$ $(n = 7l)$, (c) Martin-Kedem's method, (d) Hurd-Gerr's method ($\lfloor 1024/146 \rceil = 7$).

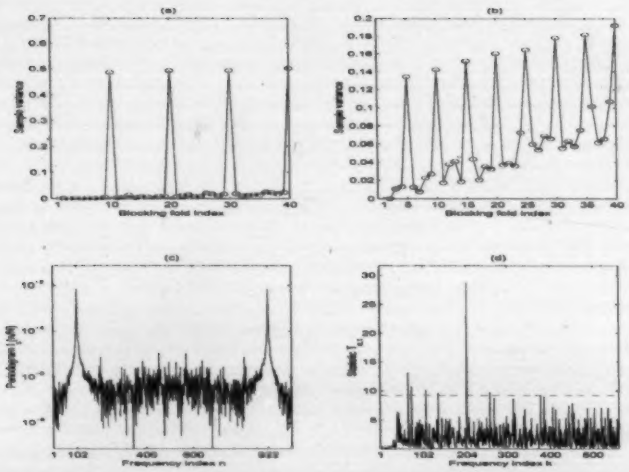

Figure 3.5: Cyclo-period inconsistency (a) The variability method based on $m_x(t)$ ($n = 10l$), (b) The variability method based on $R_{xx}(t;0)$ ($n = 5l$), (c) Martin-Kedem's method ($\lfloor 1024/102 \rfloor = 10$), (d) Dandawaté-Giannakis's method ($\lfloor 1024/204 \rfloor = 5$).

Martin and Kedem [90] solved the cyclo-period inconsistency problem (see Section 2.2) by finding the least common multiple of periods of mean and variance functions (see Section 3.5). The same idea is adopted here by simultaneously applying the variability method on the time-varying mean and the time-varying correlation.

Example 3.4 This is an example in [90] with

$$x(t) = \cos\left(\frac{2\pi t}{10}\right) + w(t),$$

where w is generated by filtering a WN$(0,1)$ signal through an AR filter with parameter 0.3. x can be shown to be both first- and second-order cyclo-stationary with periods 10 and 5, respectively (see also Lemma 2.1 in [90]). A simulation using Martin-Kedem's method gives the correct cyclo-period $p = 10$, shown in Figure 3.5-(c). Figure 3.5-(a) and (b) are the variability methods based on $m_x(t)$ and $R_x(t;0)$, respectively; the former estimates $\hat{p}_1 = 10$ and the latter gives $\hat{p}_2 = 5$; thus, their least common multiple is the correct cyclo-period $p = 10$. Hurd-Gerr's and Dandawaté-Giannakis's methods are exclusively based on some second-order statistics; therefore, they give half of the cyclo-period, $\hat{p} = 5$. Figure 3.5-(d) shows the result of Dandawaté-Giannakis's method. □

Remark: The signals like x in Example 3.4 are also known as sinusoidal or harmonic signals; thus, the cyclo-period p can be estimated by some well-known frequency estima-

tion techniques such as the Pisarenko harmonic decomposition, the MUSIC and ESPRIT algorithms (see e.g., Section 9.6 in [89]).

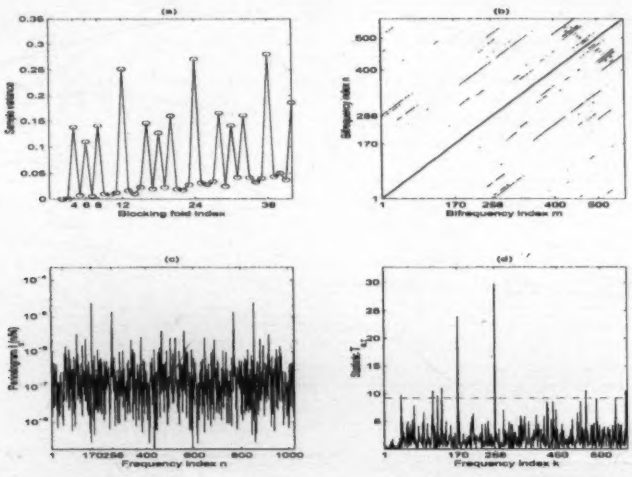

Figure 3.6: Two cyclo-stationarities (a) The variability method ($n = 12l$), (b) Hurd-Gerr's method ($\lfloor 1024/256 \rfloor = 4$, $\lfloor 1024/170 \rfloor = 6$), (c) Martin-Kedem's method ($\lfloor 1024/256 \rfloor = 4$, $\lfloor 1024/170 \rfloor = 6$), (d) Dandawaţé-Giannakis's method ($\lfloor 1024/256 \rfloor = 4$, $\lfloor 1024/170 \rfloor = 6$).

Sometimes, two (or more) cyclo-stationary signals appear at the same time[8]. If both are of interest, the cyclo-period of the integrated signal is the least common multiple of individual cyclo-periods and the detection of multiple cyclo-stationarities may be meaningful as well. On the other hand, if one of them is interpreted as noise, all the four methods are strongly affected by the noise.

Example 3.5 This is a modified version of an example in [37] (see Eq. (71) therein and Example 17.3 in [61]):

$$x(t) = \cos\left(\frac{2\pi t}{8}\right) w_1(t) + \cos\left(\frac{2\pi t}{12}\right) w_2(t),$$

where w_1 and w_2 are WN$(0,1)$ and mutually independent. x has two second-order cyclo-stationary components, one with cyclo-period 4 and the other with 6. If both components are of interest, the cyclo-period of x is the least common multiple of 4 and 6, i.e., $p = 12$. In terms of cyclo-period estimation, all methods perform well, as shown in Figure 3.6. However, the variability method cannot tell the existence of two cyclo-stationarities since the peaks at 4 and 6 may arise from a single $(\mathrm{CS})_{12}$ signal, as implied by Theorem 3.2. On the contrary, the other three methods may be capable of detecting the two cyclo-stationarities. If one of

[8]Such signals are also named as polycyclo-stationary signals [54, 56]

the two cyclo-stationary components is regarded as noise, it is clear that all methods are sensitive to it: the signal of interest may be dominated by the cyclo-stationary noise if the signal-to-noise ratio is low. □

Finally, a real-life example is presented to further testify performances of the four cyclo-period estimation methods.

Example 3.6 The global irradiance was measured hourly for three years (1990-1992) at a meteorological station DELTA on Ellsmere Island, N.W.T., Canada. Data and more information are available at the Taconite Inlet Project website[9]. As the global irradiance is one of the measurements of the solar radiation, it is plausible to conjecture that some statistic of the global irradiance has a 24 hour period rhythm; in other words, the global irradiance can be modeled as a cyclo-stationary signal with period 24 (hours). The four cyclo-period estimation methods are applied with results shown in Figure 3.7. The available data length is 4196 and the FFT lengths are chosen to be the integer in the form of 2^J ($J \in \mathbb{Z}_+$) closest to 4196, namely, 4096; while $n_{max} = 60$ and the rest parameters are the same as those used to generate Figures 3.1 and 3.2. The variability method displays a very clear periodic pattern giving $\hat{p} = 24$ that is consistent with the conjecture. However, the other three methods more or less suffer from some disturbing lines/peaks. □

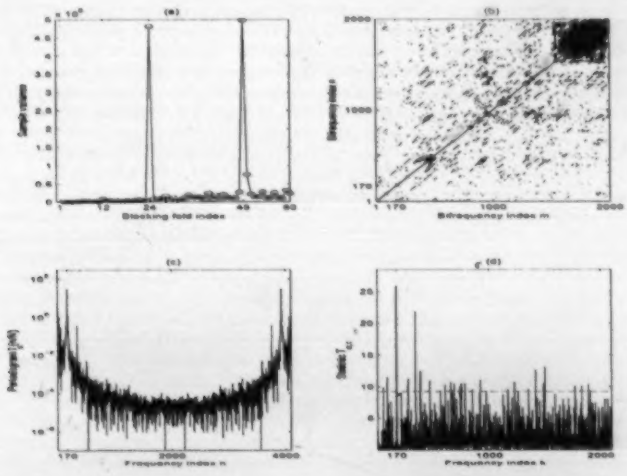

Figure 3.7: Global irradiance (a) The variability method ($n = 24l$), (b) Hurd-Gerr's method ($\lfloor 4096/170 \rfloor = 24$), (c) Martin-Kedem's method ($\lfloor 4096/170 \rfloor = 24$), (d) Dandawaté-Giannakis's method ($\lfloor 4096/170 \rceil = 24$).

[9] http://www.geo.umass.edu/climate/TILPHTML/TILPhome.html

3.3.2 Properties of the Variability Method

In terms of advantages, the variability method has the following properties:

- *Insensitive to stationary noises:* In the variability method, a stationary noise equally contributes to $\hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0)$ and $\hat{m}_{x_n^{(k)}}$ for a fixed n (consider a stationary signal as an ill-cyclo-stationary signal and apply Theorem 3.1 or Theorem 3.3); hence, the stationary noise has no effect on the variability of these statistics. Dandawaté-Giannakis's method has a noise effect cancelation in the inversion operation in (3.31). In addition, Dandawaté-Giannakis's and Hurd-Gerr's methods have thresholds to further eliminate the noise effects, but their performances are quite different in the examples. Martin-Kedem's method has a similar but somewhat incomplete cancelation in (3.26). See Examples 3.2 and 3.6.

- *Applicable to ill-cyclo-stationary signals:* Ill-cyclo-stationarity has different types in terms of $R_{xx}(t;\tau)$ (Definition 3.1) and $m_x(t)$ (Definition 3.2). According to the type of ill-cyclo-stationarity, the variability method can extract the cyclo-period from the corresponding time-varying statistics. Martin-Kedem's method is based on the mean and variance and thus is unable to deal with ill-cyclo-stationarity in $R_{xx}(t;0)$. See Example 3.3.

- *Capable of dealing with the cyclo-period inconsistency problem:* As explained in Example 3.4, the variability method and Martin-Kedem's method are able to solve the cyclo-period inconsistency problem, while Hurd-Gerr's and Dandawaté-Giannakis's methods cannot, because they are exclusively based on some second-order statistics.

- *With the best resolution:* Dandawaté and Giannakis [37] pointed out that (3.30) was usually computed via FFT, which would limit the resolution of the cyclo-period estimation. For instance, the cyclo-period p in Example 3.3 is an odd number $p = 7$, while a typical FFT length T is 2^J ($J \in \mathbb{Z}_+$); thus, p may not be estimated precisely from (3.32), i.e., $\hat{p} = \lfloor 2^J/k \rceil$, where $k \in \mathbb{Z}_+$. Hurd-Gerr's and Martin-Kedem's methods have the same problem if (3.22) and (3.27) are implemented using FFT. On the contrary, the variability method does not require FFT and provides the best resolution by searching the cyclo-period directly within the integer set.

- *Easier to use:* Besides the FFT length, other parameters have to be tuned carefully in Hurd-Gerr's, Martin-Kedem's and Dandawaté-Giannakis's methods, e.g., parameters of spectral windows, because they highly affect performances of these methods; the

effect is essentially due to the bias and variance tradeoff of the estimators used in the three methods (see e.g., Theorem 3.5 in [36]). On the contrary, the variability method has only one parameter to concern, namely, either the data length N or the largest blocking operator index n_{max}. In terms of computational complexity[10], if the correlations in (3.4) and (3.13) are computed via FFT [128], then the implementation of the variability method approximately involves $3N \log_2 \lfloor N/n \rfloor$ operations for $n = 2, 3, \cdots, n_{max}$ (see Algorithm in Section 3.3.1). Loosely speaking, the variability method performs the correlation computation (or equivalently three FFT computations) on N data points for n_{max} times; Martin-Kedem's and Dandawaté-Giannakis's methods require one and several N-point FFT computations, respectively (see Section 3.5); Hurd-Gerr's method is usually more computational expensive than the others, as (3.23) needs the M-point FFT computation for T^2 times. The execution time in one trial (Pentium 4, CPU 2.80GHz, RAM 768MB) shown in Figure 3.1 gives a rough idea on this aspect: the variability method — 1.172 sec, Hurd-Gerr's method — 39.579 sec, Martin-Kedem's method — 0.312 sec, and Dandawaté-Giannakis's method — 0.312 sec. Certainly, more efficient algorithms exist to reduce the computation costs of the four methods, e.g., the correlation algorithm at Section 5.2 in [89].

In terms of drawbacks, the variability method has the following properties:

- *Incapable of detecting the existence of two or more cyclo-stationary signals:* Hurd-Gerr's method essentially estimates and displays the bispectrum defined in (3.20), from which multiple cyclo-stationarities can be detected. Martin-Kedem's and Dandawaté-Giannakis's methods more or less have similar mechanisms. In the variability method, the cyclo-period is obtained from the variability of the time-varying correlation $R_{xx}(t;\tau)$ and/or the time-varying mean $m_x(t)$, not from these two statistics directly; when $R_{xx}(t;\tau)$ and/or $m_x(t)$ are collapsed into the variability measurements, some information is lost. See Example 3.5.

- *Reliant on a relatively large data length if the cyclo-period is large:* In the variability method, $\hat{R}_{x_n^{(k)} x_n^{(k+\tau)}}(0)$ or equivalently $\hat{R}_{xx}(t;\tau)$ (see Theorem 3.1) and $\hat{m}_{x_n^{(k)}}$ or equivalently $\hat{m}_x(t)$ (see Theorem 3.3) may be unreliable because of insufficient data points. Hence, if the cyclo-period and consequently n_{max} are large, N has to be large enough to guarantee a good estimation.

[10]The computational complexity is usually expressed by the number of complex multiplications and additions required or, simply, by the number of operations [34].

Remark: The accuracy of the variability method is essentially determined by the two estimators $\hat{R}_{xx}(t;\tau)$ and $\hat{m}_x(t)$, which are asymptotically unbiased and consistent. Theorems 3.1 and 3.3 imply that the two estimators $\hat{R}_{xx}(t;\tau)$ and $\hat{m}_x(t)$ converge to $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)$ and $\hat{m}_{x_n^{(k)}}$, respectively, in a rate of $1/N$ (see Remark after Theorem 3.1); thus, the rate of convergence for the variability method is mainly determined by those of $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)$ and $\hat{m}_{x_n^{(k)}}$, i.e., $1/\sqrt{N}$ (note that $\hat{R}_{x_n^{(k)}x_n^{(k+\tau)}}(0)$ and $\hat{m}_{x_n^{(k)}}$ are the well-known correlation and mean estimators defined for stationary signals; see Section 3.2.).

3.4 Conclusion

A new method, named as the time-varying mean/correlation variability method, is proposed to estimate the cyclo-period p of a discrete-time cyclo-stationary signal x. If x is blocked by the blocking operator L_n, $\underline{x}_n = L_n x$ defined in (2.8), estimators of the time-varying correlation and the time-varying mean are associated with some statistics of $\underline{x}_n$ in Theorems 3.1 and 3.3, respectively. When n is an integer multiple of p, Theorems 3.2 and 3.4 show that variability of these statistics of $\underline{x}_n$ achieves the maximum, from which the cyclo-period p is obtained. A detailed algorithm is given in Section 3.3.1, where simulation and real-life examples confirm the effectiveness of the variability method. The advantages and drawbacks of the variability method are summarized with the underlying rationales in Section 3.3.2.

3.5 Appendix

The purpose of this appendix is two-fold: (i) main ideas of Hurd-Gerr's, Martin-Kedem's, and Dandawaté-Giannakis's methods are introduced in detail to prepare a comparison of them with the variability method; (ii) it is meaningful to present algorithms of the three methods, since Hurd and Gerr [72] and Martin and Kedem [90] did not provide such detailed algorithms and Dandawaté and Giannakis [37] had one in a broader context. Before introducing the three methods, three concepts are reviewed, namely, the cyclic covariance, the cyclic spectrum and the bispectrum.

Cyclic Covariance and Cyclic Spectrum: As $R_{xx}(t;\tau)$ in (2.6) is a periodic sequence of t with period p for a fixed τ, it has a discrete Fourier series expansion for $\alpha = 2\pi k/p$,

$$C_{xx}(\alpha;\tau) = \frac{1}{p}\sum_{t=0}^{p-1} R_{xx}(t;\tau)\, e^{-j\alpha t}. \tag{3.18}$$

For a fixed α, $C_{xx}(\alpha;\tau)$, called the cyclic covariance [62, 56, 37], is a sequence in τ; thus,

it has a DTFT with respect to τ,

$$S_{xx}(\alpha;\omega) = \sum_{\tau=-\infty}^{\infty} C_{xx}(\alpha;\tau)\, e^{-j\omega\tau}, \tag{3.19}$$

which is defined as the cyclic spectrum of x [56].

Bispectrum: The bispectrum[11] $S_{xx}(v,\omega)$ is defined as the two-dimensional DTFT of the correlation $R_{xx}(t_1,t_2) = E\{x(t_1)x^*(t_2)\}$ [100, 2],

$$S_{xx}(v,\omega) = \frac{1}{2\pi} \sum_{t_1=-\infty}^{\infty} \sum_{t_2=-\infty}^{\infty} R_{xx}(t_1,t_2)\, e^{-jvt_1} e^{j\omega t_2}. \tag{3.20}$$

The bispectrum is a very general concept describing the second-order statistical property of a non-stationary signal. Specifically, the bispectrum of a cyclo-stationary signal lies on some parallel lines in the v-ω plane [2, 62],

$$\omega - v + \frac{2\pi k}{p} = 0, \quad k \in \mathbb{Z}. \tag{3.21}$$

The bispectrum component on the k-th line is equal to the cyclic spectrum $S_{xx}(2\pi k/p;\omega)$ in (3.19) [112, 2].

Hurd-Gerr's Method

Hurd and Gerr [72] developed their method by observing that for a cyclo-stationary signal x, the discrete Fourier transformation (DFT) coefficients of a finite length sample of x exhibit (spectral) correlation in disjoint frequency bands described in (3.21). Alternatively, a more transparent interpretation is the following: it is well-known that the power spectrum of a stationary signal can be estimated by a nonparametric method, the periodogram (see e.g., [68, 89]); analogously, the bispectrum defined in (3.20) can be estimated by a two-dimensional periodogram; the estimated bispectra lie on the parallel lines represented in (3.21), from which the cyclo-period is obtained.

Algorithm: Given $\{x(t)\}_{t=0}^{N-1}$, the cyclo-period p can be estimated by the following steps:

1. Compute the DFT of $\{x(t)\}_{t=0}^{N-1}$ (usually implemented by FFT),

$$X_T(\omega_k) = \sum_{t=0}^{T-1} x(t)\, e^{-j\omega_k t}, \tag{3.22}$$

where T is a selected DFT length and $\omega_k = 2\pi k/T$.

[11] The terminology "bispectrum" has a different meaning in the literature, namely, the two-dimensional DTFT of the *third*-order cumulant [135].

2. Select a positive integer $M \ll N$ and compute a normalized[12] two-dimensional periodogram

$$|\gamma(m,n,M)|^2 = \frac{\left|\sum_{k=0}^{M-1} X_T(\omega_{m+k})\, X_T^*(\omega_{n+k})\right|^2}{\sum_{k=0}^{M-1} |X_T(\omega_{m+k})|^2 \sum_{k=0}^{M-1} |X_T(\omega_{n+k})|^2}, \tag{3.23}$$

for $m, n = 0, 1, \cdots, T-1$.

3. Pick up a confidence level P_0, e.g., 99% or 95%, and calculate a threshold $|\gamma_0|^2$ from

$$\Pr\left(|\gamma(m,n,M)|^2 > |\gamma_0|^2\right) = \left(1 - |\gamma_0|^2\right)^{M-1} = 1 - P_0. \tag{3.24}$$

4. Map $|\gamma(m,n,M)|^2$ that exceeds $|\gamma_0|^2$ proportionally to its magnitude on the m-n plane[13], from which the cyclo-period is estimated as

$$\hat{p} = \lfloor T/d \rceil,$$

where d is the minimum difference between m and n giving off-diagonal bispectrum lines.

Martin-Kedem's Method

Martin-Kedem's method was mainly motived by a specific case of the cyclo-period inconsistency problem introduced in Section 2.2, namely, the period of mean function is p while that of variance function is $p/2$ [90]. To deal with this problem, a new sequence is formed from the original data such that the new sequence is periodic with the same period as the least common multiple of the periods of mean and variance functions of the original cyclo-stationary signal; next, the periodogram of the new sequence displays peaks at certain frequencies, from which the cyclo-period can be estimated.

Algorithm: Given $\{x(t)\}_{t=0}^{N-1}$, the cyclo-period p can be estimated by the following steps:

1. Choose an integer K so that x is K-dependent in terms of covariance, i.e.,

$$E\{[x(t) - m_x(t)]\,[x(t+k) - m_x(t+k)]\} = 0, \quad \forall k > K. \tag{3.25}$$

2. Select an integer M such that $K < M < N$, and form a new sequence s,

$$s(k) = \frac{1}{N-k} \sum_{t=0}^{N-k-1} |x(t) - x(t+k)|, \quad K < k \le M. \tag{3.26}$$

[12] Eq. (3.23) implies that $0 \le |\gamma(m,n,M)|^2 \le 1$ and $|\gamma(m,n,M)|^2 = 1$ if $m = n$ and x is real-valued.
[13] The mapping can be implemented, e.g., by Matlab function "contour".

3. Form a shifted version of s, $s_o(k) = s(k+K)$ and compute the periodogram of s_o for a selected DFT length T,

$$I_o(n) = \frac{2}{T}\left|\sum_{k=1}^{T} s_o(k)\, e^{-j2\pi kn/T}\right|^2, \text{for } n = 1, 2, \cdots, T. \tag{3.27}$$

4. Locate the smallest n giving the highest peak of $I_o(n)$ and estimate the cyclo-period

$$\hat{p} = \lfloor T/n \rceil .$$

Dandawaté-Giannakis's Method

The main idea of Dandawaté-Giannakis's method is to consistently estimate $C_{xx}(\alpha;\tau)$ in (3.18) that is zero except at $\alpha = 2\pi k/p$ ($k \in \mathbb{Z}$); thus, α and equivalently p can be obtained by checking if $C_{xx}(\alpha;\tau)$ is zero from a statistical χ^2 test, which is based on the asymptotic distribution of estimation error and a consistent estimator of $S_{xx}(\alpha;\omega)$ defined in (3.19). Note that both time- and frequency-domain methods were given in [37], whereas only the time-domain method is presented here because it is more computationally convenient in the context of cyclo-stationary signals [37].

Algorithm: Given $\{x(t)\}_{t=0}^{N-1}$ and a fixed delay τ, the cyclo-period p can be estimated by the following steps:

1. Compute a row vector for some integer $k \in [1, T]$ (T is a selected DFT length),

$$\hat{c}_{k,\tau} = \left[\ \mathrm{Re}\left\{\tfrac{1}{T}\textstyle\sum_{t=1}^{T} x(t)\,x(t+\tau)\,e^{-\frac{j2\pi kt}{T}}\right\}\quad \mathrm{Im}\left\{\tfrac{1}{T}\textstyle\sum_{t=1}^{T} x(t)\,x(t+\tau)\,e^{-\frac{j2\pi kt}{T}}\right\}\ \right],$$

where $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ represent the real and imaginary parts, respectively.

2. Estimate an error covariance matrix

$$\hat{\Phi}_{k,\tau} = \begin{bmatrix} \mathrm{Re}\left\{\frac{\hat{Q}_{k,\tau}+\hat{Q}^{(*)}_{k,\tau}}{2}\right\} & \mathrm{Im}\left\{\frac{\hat{Q}_{k,\tau}-\hat{Q}^{(*)}_{k,\tau}}{2}\right\} \\ \mathrm{Im}\left\{\frac{\hat{Q}_{k,\tau}+\hat{Q}^{(*)}_{k,\tau}}{2}\right\} & \mathrm{Re}\left\{\frac{\hat{Q}^{(*)}_{k,\tau}-\hat{Q}_{k,\tau}}{2}\right\} \end{bmatrix},$$

where

$$\hat{Q}_{k,\tau} = \frac{1}{TL}\sum_{t=-(L-1)/2}^{(L-1)/2} W_L(t)\, F_\tau\left(\frac{2\pi(k-t)}{T}\right) F_\tau\left(\frac{2\pi(k+t)}{T}\right), \tag{3.28}$$

$$\hat{Q}^{(*)}_{k,\tau} = \frac{1}{TL}\sum_{t=-(L-1)/2}^{(L-1)/2} W_L(t)\, F^*_\tau\left(\frac{2\pi(k+t)}{T}\right) F_\tau\left(\frac{2\pi(k+t)}{T}\right). \tag{3.29}$$

Here W_L is a smoothing window with an odd length L, e.g., Kaiser window (see e.g. [68, 104]), and

$$F_\tau(\omega) = \sum_{t=1}^{T} x(t)\, x(t+\tau)\, e^{-j\omega t}. \tag{3.30}$$

3. Compute a real-valued test statistic

$$\Gamma_{k,\tau} = T\hat{c}_{k,\tau}\hat{\Phi}_{k,\tau}^{-1}(\hat{c}_{k,\tau})^*, \tag{3.31}$$

choose a false alarm level P_F, e.g., 1% or 5%, and find a threshold Γ from the χ^2 table with 2 degrees of freedom so that $\Pr\{\Gamma_{k,\tau} > \Gamma\} = P_F$ under a hypothesis $C_{xx}(2\pi k/T;\tau) = 0$.

4. Repeat the above steps with different k's and find the smallest k such that $\Gamma_{k,\tau} > \Gamma$; the cyclo-period is estimated as

$$\hat{p} = \lfloor T/k \rceil. \tag{3.32}$$

Chapter 4

Cyclo-Statistic Estimation

This chapter[1] studies the first- and second-order statistic estimators for discrete-time cyclo-stationary signals, with a focus on those of practical interests — the estimators of the time-varying mean/correlation and the cyclic correlation/spectrum. A new cyclic spectrum estimator, based on the blocked representation of cyclo-stationary signals, is proposed. The rationale of an implementation shortcut for the cyclic mean/correlation/spectrum estimator is explored from the relationship between cyclo-stationarity and quasi-stationarity. Performance of the cyclo-statistic estimators is validated via simulation examples.

4.1 Introduction

A new cycle domain appears in cyclo-stationary signals due to the time-varying features of mean and correlation functions; thus, cyclo-statistics are much richer than those of stationary signals. Among all cyclo-statistics, the time-varying mean/correlation and the cyclic correlation/spectrum are of practical interests (to be discussed later). The contribution of this chapter, which also has some tutorial value, is to summarize the existing estimators of these cyclo-statistics, to supplement a new blocking-based cyclic spectrum estimator, and to discuss implementation issues of these estimators.

Some conditions being generalized from those for stationary signals are assumed to hold in order to consistently estimate cyclo-statistics from one sample realization, e.g., cyclo-ergodicities [24] and finite *memories* of the time-varying correlation and the cyclic correlation [36]. Along with cyclo-ergodicities, Boyles and Gardner [24] presented estimators of the time-varying mean and the cyclic mean, which also appeared in bits and pieces in other work, e.g., [37, 94]. By dropping lim$_{\cdot}$ and expectations from the corresponding definitions, estimators of the cyclic correlation and the cyclic spectrum were proposed and

[1]The chapter has been published in [155].

analyzed from various perspectives. Dandawaté and Giannakis [36] defined the higher-order cyclic auto-cumulants/spectra and developed an asymptotic theory for the higher-order cyclic auto-spectra in a broader context of almost cyclo-stationary signals; they further completed the asymptotic theory for the higher-order cyclic auto-cumulants in [38]. Later, Sadler and Dandawaté [111] generalized the asymptotic theory to the cyclic cross-spectrum between jointly cyclo-stationary signals. Genossar *et al.* [59] investigated conditions to guarantee consistency of the cyclic auto-correlation estimator. Schell [115] explored asymptotic moments of the cyclic cross-correlation estimators for multivariate cyclo-stationary signals. Alternatively, the cyclic correlation/spectrum and the time-varying correlation/spectrum can be estimated by taking advantage of the inherent periodicity. Alekseev [3] estimated the cyclic spectrum starting from a time-varying correlation estimator, which was the same as the synchronous average in [94]. Sakai and Ohno [112] pointed out a possibility of estimating the cyclic spectrum from spectra of blocked signals, which is formally proposed as the blocking-based estimator of the cyclic spectrum in this chapter (see Section 4.4). In terms of implementation, Gardner [56] offered an observation, which has been ignored in the literature, that the cyclic auto-spectrum is equivalent to a particular cross-spectrum defined for stationary signals. In fact, the observation is applicable for a class of cyclo-statistic estimators; its underlying rationale is based on the relationship between cyclo-stationarity and quasi-stationarity (see Section 4.5). Note that some of the above estimators have their counterparts for continuous-time cyclo-stationary signals [52, 73, 25, 41].

The rest of the chapter is organized as follows. Cyclo-statistic estimators are studied in Sections 4.2, 4.3 and 4.4, on the cyclo-mean, correlation and spectrum, respectively. Section 4.5 explores the relationship between cyclo-stationarity and quasi-stationarity in order to theoretically support an implementation shortcut of a class of cyclo-statistic estimators. Finally, Section 4.6 provides concluding remarks.

4.2 Cyclo-Mean Estimators

Estimators of the time-varying mean and the cyclic mean are summarized in this section with an emphasis on the former for its practical usages: before estimating the second-order cyclo-statistics, the time-varying mean usually has to be subtracted from data in order to avoid spectral leakages around low frequencies — similar to a common data processing step for stationary signals (see e.g., [89, 119]); moreover, valuable information, e.g., the cyclo-period p can be extracted from the time-varying mean (see Chapter 3).

The time-varying mean $m_x(t)$ of a $(\mathrm{CS})_p$ signal $x(t)$ has a Discrete Fourier Series (DFS)

expansion because of the periodicity in (2.5), i.e.,

$$m_x(t) = \sum_{k=0}^{p-1} M_x(k)\, e^{j2\pi kt/p}, \tag{4.1}$$

where

$$M_x(k) = \frac{1}{p}\sum_{t=0}^{p-1} m_x(t)\, e^{-j2\pi kt/p}. \tag{4.2}$$

Here k and $M_x(k)$ are usually named as the cycle index and the cyclic mean, respectively [56, 61]. Eq. (4.2) implies that $M_x(k)$, like $m_x(t)$, is a periodic sequence with period p.

Given one sample realization of $x(t)$,

$$\{x(t)\}_{t=0}^{N-1} := \{x(0), x(1), \cdots, x(N-1)\}, \tag{4.3}$$

the first estimator of $m_x(t)$ is based on the periodicity in (2.5) [24, 94],

$$\hat{m}_x^{(i)}(t) = \frac{1}{L}\sum_{l=0}^{L-1} x(lp+t), \tag{4.4}$$

where $L = \lfloor N/p \rfloor$[2] and $t \in [0, p-1]$. The second estimator is from the definitions in (4.1) and (4.2) [24, 37, 94],

$$\hat{m}_x^{(ii)}(t) = \sum_{k=0}^{p-1} \hat{M}_x(k)\, e^{j2\pi kt/p}, \tag{4.5}$$

where $\hat{M}_x(k)$ is a cyclic mean estimator,

$$\hat{M}_x(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)\, e^{-j2\pi kn/p}. \tag{4.6}$$

$\hat{m}_x^{(i)}(t)$ and $\hat{m}_x^{(ii)}(t)$ are connected:

$$\begin{aligned}
\hat{m}_x^{(ii)}(t) &= \sum_{k=0}^{p-1}\left(\frac{1}{N}\sum_{n=0}^{N-1} x(n)\, e^{-j2\pi kn/p}\right) e^{j2\pi kt/p} \\
&\simeq \frac{1}{N}\sum_{k=0}^{p-1}\sum_{l=0}^{\lfloor N/p\rfloor-1}\sum_{r=0}^{p-1} x(lp+r)\, e^{-j2\pi k(lp+r)/p} e^{j2\pi kt/p} \\
&= \frac{1}{N/p}\sum_{l=0}^{\lfloor N/p\rfloor-1}\sum_{r=0}^{p-1} x(lp+r)\,\frac{1}{p}\sum_{k=0}^{p-1} e^{-j2\pi k(r-t)/p} \\
&= \frac{1}{N/p}\sum_{l=0}^{\lfloor N/p\rfloor-1} x(lp+t),
\end{aligned} \tag{4.7}$$

[2]Data points $x(p\lfloor N/p\rfloor)$, $x(p\lfloor N/p\rfloor+1)$, $\cdots$, $x(N)$ may be discarded in the blocking operation.

where the last equality is from an identity

$$\frac{1}{p}\sum_{k=0}^{p-1} e^{-j2\pi k(r-t)/p} = \begin{cases} 1, & (r-t)/p \text{ is an integer,} \\ 0, & \text{otherwise.} \end{cases} \tag{4.8}$$

Therefore, if $\lfloor N/p \rfloor \equiv N/p$, $\hat{m}_x^{(i)}(t)$ and $\hat{m}_x^{(ii)}(t)$ are the same; otherwise, they have a minor difference that is negligible for a large N.

Properties of $\hat{m}_x^{(i)}(t)$ can be obtained from the blocked representation of $x(t)$, shown as follows. A new signal $u_t(l) := x(lp+t)$ in (4.4) for a fixed t is the same as the t-th component of $\underline{x}_p(l)$ in (2.8) that is stationary (see Section 2.3). Thus, $\hat{m}_x^{(i)}(t)$ is the same as the familiar mean estimator defined for stationary signals and is unbiased and consistent. In particular, the variance of $\hat{m}_x^{(i)}(t)$ is obtained from the variance expression of the mean estimator of $u_t(l)$ (see e.g., Section 3.6.2 in [89]),

$$\mathrm{Var}\left\{\hat{m}_x^{(i)}(t)\right\} = \frac{\sigma_{u_t}^2}{L}\left[1 + 2\sum_{l=0}^{L-1}\left(1-\frac{l}{L}\right)\frac{\gamma_{u_t u_t}(l)}{\sigma_{u_t}^2}\right], \tag{4.9}$$

where $\gamma_{u_t u_t}(\cdot)$ and $\sigma_{u_t}^2$ are the covariance sequence and the variance of $u_t(l)$, respectively.

Example 4.1 A $(\mathrm{CS})_4$ signal is

$$x(t) = \left(1 + \cos\left(\frac{2\pi t}{4}\right)\right)(1 + e(t)),$$

where $e(t)$ is a white noise with zero-mean and unit-variance, abbreviated as WN(0, 1). Table 4.1 presents the true time-varying mean $m_x(t)$ and its estimates $\hat{m}_x^{(i)}(t)$ and $\hat{m}_x^{(ii)}(t)$ from one typical Monte Carlo trial. In addition, $\hat{m}_x^{(i)}(t)$ is given with the 95% confidence level that is developed from (4.9) and the central limit theorem. The simulation results show that $m_x(t)$ is effectively estimated by $\hat{m}_x^{(i)}(t)$ and $\hat{m}_x^{(ii)}(t)$; $\hat{m}_x^{(i)}(t)$ and $\hat{m}_x^{(ii)}(t)$ are almost the same ($N = 4001$ and $p = 4$); the variance of $\hat{m}_x^{(i)}(t)$ is validated by the reasonable confidence level. □

$N = 4001$	$t = 0$	$t = 1$	$t = 2$	$t = 3$
$m_x(t)$	2	1	0	1
$\hat{m}_x^{(i)}(t)$	2.0561 ± 0.0714	0.9852 ± 0.0152	0 ± 0	0.9834 ± 0.0158
$\hat{m}_x^{(ii)}(t)$	2.0556	0.9853	0	0.9832

Table 4.1: Estimate the time-varying mean

4.3 Cyclo-Correlation Estimators

This section studies estimators of the time-varying correlation and the cyclic correlation for jointly cyclo-stationary signals. These estimators are the bases of the cyclic correlogram (see Section 4.4); more importantly, they have some direct applications, e.g., cyclic correlation analysis based on the cyclic correlation (Chapter 6), and cyclo-period estimation based on the time-varying correlation (Chapter 3) and on the cyclic correlation [37].

The time-varying correlation $r_{xy}(t;\tau) := E\{x(t+\tau)y^*(t)\}$ of two jointly cyclo-stationary signals $x(t)$ and $y(t)$ is periodic in t for any fixed τ, i.e.,

$$r_{xy}(t+lp;\tau) = r_{xy}(t;\tau), \forall t, l \in \mathbb{Z}. \tag{4.10}$$

The periodicity leads to a DFS expansion,

$$r_{xy}(t;\tau) = \sum_{k=0}^{p-1} R_{xy}(k;\tau)\, e^{j2\pi kt/p}, \tag{4.11}$$

where

$$R_{xy}(k;\tau) = \frac{1}{p}\sum_{t=0}^{p-1} r_{xy}(t;\tau)\, e^{-j2\pi kt/p}. \tag{4.12}$$

Here $R_{xy}(k;\tau)$ is usually called the cyclic correlation [62, 56]. Eq. (4.12) implies that $R_{xy}(k;\tau)$ is periodic in k with period p.

Given $\{x(t)\}_{t=0}^{N-1}$ and $\{y(t)\}_{t=0}^{N-1}$ as that in (4.3), the first estimator of $r_{xy}(t;\tau)$ is from the periodicity in (4.10) [3, 94, 61],

$$\hat{r}_{xy}^{(i)}(t;\tau) = \frac{1}{L_1}\sum_{l=0}^{L_2-1} x(lp+t+\tau)\, y^*(lp+t), \tag{4.13}$$

where $L_1 = \lfloor N/p \rfloor$, $L_2 = \lfloor (N-\tau)/p \rfloor$, $t \in [0, p-1]$ and $\tau \in [0, N-1]$. Alternatively, $r_{xy}(t;\tau)$ can be estimated based on the definitions in (4.11) and (4.12) [36, 38, 59],

$$\hat{r}_{xy}^{(ii)}(t;\tau) = \sum_{k=0}^{p-1} \hat{R}_{xy}(k;\tau)\, e^{j2\pi kt/p}, \tag{4.14}$$

where $\tau \in [0, N-1]$ and $\hat{R}_{xy}(k;\tau)$ is a cyclic correlation estimator,

$$\hat{R}_{xy}(k;\tau) = \frac{1}{N}\sum_{n=0}^{N-\tau-1} x(n+\tau)\, y^*(n)\, e^{-j2\pi kn/p}. \tag{4.15}$$

For a negative time difference τ, the above estimators work with the periodicities of $r_{xy}(t;\tau)$ and $R_{xy}(k;\tau)$ and the following properties:

$$r_{xy}(t;-\tau) = r_{yx}^*(t-\tau;\tau), \quad R_{xy}(k;-\tau) = R_{yx}^*(-k;\tau)\, e^{-j2\pi k\tau/p}, \quad \forall t, k, l \in \mathbb{Z}.$$

It can be shown via a derivation similar to (4.7) that $\hat{r}_{xy}^{(i)}(t;\tau)$ and $\hat{r}_{xy}^{(ii)}(t;\tau)$ are the same for $N = lp$ ($l \in \mathbb{Z}_+$) or almost the same otherwise.

The estimators $\hat{r}_{xy}^{(i)}(t;\tau)$, $\hat{r}_{xy}^{(ii)}(t;\tau)$ and $\hat{R}_{xy}(k;\tau)$ are asymptotically unbiased and consistent. The mean and variance of $\hat{R}_{xy}(k;\tau)$ were given in [59, 115], while those of $\hat{r}_{xy}^{(i)}(t;\tau)$ can be reached with the help of the blocked representations of $x(t)$ and $y(t)$ as follows. Let $u_{t,\tau}(l) := x(lp + t + \tau)$ and $v_t(l) := y(lp + t)$ for fixed t and τ. Revisiting (2.8) gives that $u_{t,\tau}(l)$ and $v_t(l)$ are the components of the blocked signals $\underline{x}_p(l)$ and $\underline{y}_p(l)$, respectively. Since $u_{t,\tau}(l)$ and $v_t(l)$ are jointly stationary (see Section 2.3), $\hat{r}_{xy}^{(i)}(t;\tau)$ is the same as the well-known correlation estimator of stationary signals at lag 0, i.e., $\hat{r}_{u_{t,\tau}v_t}(0)$. Thus, the mean and the variance of $\hat{r}_{xy}^{(i)}(t;\tau)$ are ready to be obtained from those of $\hat{r}_{u_{t,\tau}v_t}(0)$ (see e.g., Section 9.2 in [119]),

$$E\left\{\hat{r}_{xy}^{(i)}(t;\tau)\right\} = \frac{L_2}{L_1} r_{xy}(t;\tau)$$

and

$$\mathrm{Var}\left\{\hat{r}_{xy}^{(i)}(t;\tau)\right\} \simeq \frac{1}{L_1} \sum_{k=-\infty}^{\infty} \left[r_{u_{t,\tau}u_{t,\tau}}(k)\, r_{v_tv_t}(k) + r_{u_{t,\tau}v_t}(k)\, r_{v_tu_{t,\tau}}(k)\right]. \tag{4.16}$$

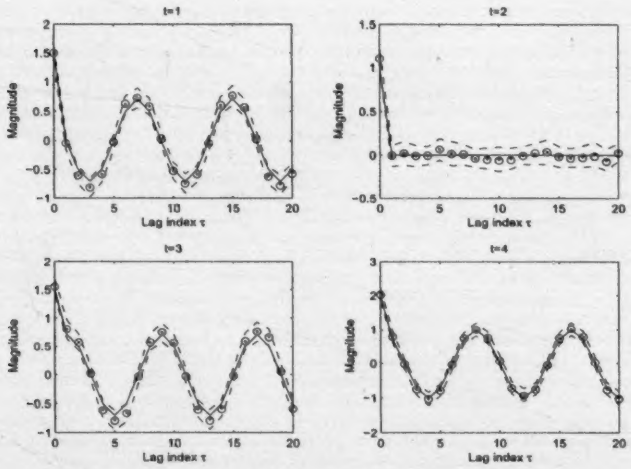

Figure 4.1: Estimate the time-varying correlation ($N = 4001$): $r_{xx}(t;\tau)$ (solid), $\hat{r}_{xx}^{(i)}(t;\tau)$ (dotts), $\hat{r}_{xx}^{(ii)}(t;\tau)$ (circles) and the 95% confidence interval of $\hat{r}_{xx}^{(i)}(t;\tau)$ (dash).

Example 4.2 This is one of a class of cyclo-stationary signals [90]:

$$x(t) = \cos\left(\frac{2\pi t}{8}\right) + e(t).$$

Here $e(t)$ is WN$(0,1)$. The time-varying correlation of $x(t)$ is

$$r_{xx}(t;\tau) = \frac{1}{2}\cos\left(\frac{2\pi t}{4} + \frac{2\pi\tau}{8}\right) + \frac{1}{2}\cos\left(\frac{2\pi\tau}{8}\right) + \delta(\tau),$$

where $\delta(\tau)$ is the discrete Delta function,

$$\delta(\tau) = \begin{cases} 1, & \tau = 0 \\ 0, & \text{elsewhere.} \end{cases}$$

Figure 4.1 presents the true time-varying correlation $r_{xx}(t;\tau)$ and the estimates $\hat{r}_{xx}^{(i)}(t;\tau)$ and $\hat{r}_{xx}^{(ii)}(t;\tau)$ in one typical Monte Carlo trial. The 95% confidence interval of $\hat{r}_{xx}^{(i)}(t;\tau)$ is obtained based on (4.16). The simulation results show that $r_{xx}(t;\tau)$ is effectively estimated by $\hat{r}_{xx}^{(i)}(t;\tau)$ and $\hat{r}_{xx}^{(ii)}(t;\tau)$; $\hat{r}_{xx}^{(i)}(t;\tau)$ and $\hat{r}_{xx}^{(ii)}(t;\tau)$ are almost the same ($N = 4001$ and $p = 4$); the variance of $\hat{r}_{xx}^{(i)}(t;\tau)$ gives a reasonable confidence interval. □

4.4 Cyclic Spectrum Estimators

Cyclo-stationary signals have at least four spectral representations that are closely related and mutually convertible (Chapter 5), namely, the cyclic spectrum, the time frequency representation, the bispectrum and the 2-D spectrum, among which the cyclic spectrum is the most common and convenient and hence is the one to be focused on. This section presents three types of cyclic spectrum estimators: the cyclic periodogram, the cyclic correlogram and the blocking-based estimator.

For jointly cyclo-stationary signals $x(t)$ and $y(t)$, their cyclic spectrum is defined as the DTFT of the cyclic correlation $R_{xy}(k;\tau)$ with respect to τ [62, 56],

$$S_{xy}\left(k;e^{j\omega}\right) = \sum_{\tau=-\infty}^{\infty} R_{xy}(k;\tau)\, e^{-j\omega\tau}. \tag{4.17}$$

Equivalently, $S_{xy}\left(k;e^{j\omega}\right)$ can be represented in terms of the time-varying correlation $r_{xy}(t;\tau)$ by inserting (4.12) into (4.17),

$$S_{xy}\left(k;e^{j\omega}\right) = \frac{1}{p}\sum_{t=0}^{p-1}\sum_{\tau=-\infty}^{\infty} r_{xy}(t;\tau)\, e^{-j2\pi kt/p} e^{-j\omega\tau}. \tag{4.18}$$

A different viewpoint defines $S_{xy}\left(k;e^{j\omega}\right)$ as the cyclic power spectral density [61],

$$S_{xy}\left(k;e^{j\omega}\right) = \lim_{N\to\infty}\frac{1}{N}E\left\{X_N\left(e^{j\omega}\right)Y_N^*\left(e^{j(\omega-2\pi k/p)}\right)\right\}. \tag{4.19}$$

Here $X_N\left(e^{j\omega}\right)$ and $Y_N\left(e^{j\omega}\right)$ respectively are the N-point DFTs of $x(t)$ and $y(t)$, e.g.,

$$X_N\left(e^{j\omega}\right) = \sum_{t=0}^{N-1} x(t)\, e^{-j\omega t}.$$

Either (4.17), (4.18) or (4.19) implies that $S_{xy}\left(k;\mathrm{e}^{j\omega}\right)$ is periodic in k with period p.

Given $\{x(t)\}_{t=0}^{N-1}$ and $\{y(t)\}_{t=0}^{N-1}$, $S_{xy}\left(k;\mathrm{e}^{j\omega}\right)$ can be estimated by the following three types of estimators.

Cyclic Periodogram: The cyclic periodogram is obtained by dropping the limit and the expectation from the definition of $S_{xy}\left(k;\mathrm{e}^{j\omega}\right)$ in (4.19),

$$\hat{S}_{xy}^{(\mathrm{i})}\left(k;\mathrm{e}^{j\omega}\right)=\frac{1}{N}X_N\left(\mathrm{e}^{j\omega}\right)Y_N^*\left(\mathrm{e}^{j(\omega-2\pi k/p)}\right). \tag{4.20}$$

Unfortunately, $\hat{S}_{xy}^{(\mathrm{i})}\left(k;\mathrm{e}^{j\omega}\right)$ is inconsistent [36, 37, 111]. To achieve the consistency, a smoothed cyclic periodogram is often used [36, 111],

$$\hat{S}_{xy}^{(L)}\left(k;\mathrm{e}^{j\omega}\right)=\frac{1}{L}\sum_{l=-(L-1)/2}^{(L-1)/2}\hat{S}_{xy}^{(\mathrm{i})}\left(k;\mathrm{e}^{j2\pi l/L}\right)\cdot W^{(L)}\left(\omega-\frac{2\pi l}{L}\right). \tag{4.21}$$

Here $W^{(L)}(\cdot)$ is a spectrum smoothing window having some properties as those in assumptions 3.1-3.6 in [36] or assumption 2 in [111]. $\hat{S}_{xy}^{(L)}\left(k;\mathrm{e}^{j\omega}\right)$ actually adopts the idea in Blackman-Tukey's estimator that estimates spectra of stationary signals (see e.g., Chapter 8 in [68] or Chapter 5 in [89]) by smoothing contiguous values of one single cyclic periodogram. As Blackman-Tukey's estimator, $\hat{S}_{xy}^{(L)}\left(k;\mathrm{e}^{j\omega}\right)$ is asymptotically unbiased and consistent [36, 111]. Note that $\hat{S}_{xy}^{(L)}\left(k;\mathrm{e}^{j\omega}\right)$ is not the only choice to make $\hat{S}_{xy}^{(\mathrm{i})}\left(k;\mathrm{e}^{j\omega}\right)$ consistent; see Section 4.5.

Cyclic Correlogram: The definitions in (4.17) and (4.18) lead to two cyclic correlograms:

$$\hat{S}_{xy}^{(\mathrm{ii}A)}\left(k;\mathrm{e}^{j\omega}\right)=\sum_{\tau=-(N-1)}^{N-1}\hat{R}_{xy}\left(k;\tau\right)\mathrm{e}^{-j\omega\tau} \tag{4.22}$$

and

$$\hat{S}_{xy}^{(\mathrm{ii}B)}\left(k;\mathrm{e}^{j\omega}\right)=\frac{1}{p}\sum_{t=0}^{p-1}\sum_{\tau=-(N-1)}^{N-1}\hat{r}_{xy}^{(\mathrm{i})}\left(t;\tau\right)\mathrm{e}^{-j2\pi kt/p}\mathrm{e}^{-j\omega\tau}. \tag{4.23}$$

Here $\hat{r}_{xy}^{(\mathrm{i})}\left(t;\tau\right)$ and $\hat{R}_{xy}\left(k;\tau\right)$ are given in (4.13) and (4.15), respectively. In practice, a good modification to (4.22) or (4.23) is to assign different weightings, say $\mathrm{w}(\tau)$, to $\hat{r}_{xy}^{(\mathrm{i})}\left(t;\tau\right)$ or $\hat{R}_{xy}\left(k;\tau\right)$ in order to penalize points with higher variances. The estimators $\hat{S}_{xy}^{(\mathrm{ii}A)}\left(k;\mathrm{e}^{j\omega}\right)$ and $\hat{S}_{xy}^{(\mathrm{ii}B)}\left(k;\mathrm{e}^{j\omega}\right)$ are the same or almost the same because of (4.14) and the closeness between $\hat{r}_{xy}^{(\mathrm{i})}\left(t;\tau\right)$ and $\hat{r}_{xy}^{(\mathrm{ii})}\left(t;\tau\right)$; in addition, they are equivalent to the smoothed cyclic periodogram $\hat{S}_{xy}^{(L)}\left(k;\mathrm{e}^{j\omega}\right)$, if the window $W^{(L)}(\cdot)$ in (4.21) is the DTFT of the weighting $\mathrm{w}(\tau)$ assigned to $\hat{r}_{xy}^{(\mathrm{i})}\left(t;\tau\right)$ or $\hat{R}_{xy}\left(k;\tau\right)$. This equivalence can be proved by generalizing the relationship between Blackman-Tukey's estimator and the correlogram defined for stationary signals (see e.g., Section 5.3.2 in [89]). Therefore, $\hat{S}_{xy}^{(\mathrm{ii}A)}\left(k;\mathrm{e}^{j\omega}\right)$ and $\hat{S}_{xy}^{(\mathrm{ii}B)}\left(k;\mathrm{e}^{j\omega}\right)$

are asymptotically unbiased and consistent, as $\hat{r}_{xy}^{(i)}(t;\tau)$, $\hat{R}_{xy}(k;\tau)$ and $\hat{S}_{xy}^{(L)}\left(k;\mathrm{e}^{j\omega}\right)$ posses these properties.

Blocking-Based Estimator: Considering a fact that the blocked signals $\underline{x}_p(l)$ and $\underline{y}_p(l)$ are jointly stationary (see Section 2.3), the cyclic spectrum can be estimated as

$$\hat{S}_{xy}^{(iii)}\left(k;\mathrm{e}^{j\omega}\right) = \left[U_p\left(\mathrm{e}^{j\omega}\right)\cdot \hat{S}_{\underline{x}_p\underline{y}_p}\left(\mathrm{e}^{jp\omega}\right)\cdot U_p^*\left(\mathrm{e}^{j\omega}\right)\right]_{0,k}. \tag{4.24}$$

Here $\hat{S}_{\underline{x}_p\underline{y}_p}\left(\mathrm{e}^{jp\omega}\right)$ is an estimate of the spectrum of $\underline{x}_p(l)$ and $\underline{y}_p(l)$[3]; $U_p\left(\mathrm{e}^{j\omega}\right)$ is a unitary matrix whose element at the k-th row and the l-th column is

$$\left[U_p\left(\mathrm{e}^{j\omega}\right)\right]_{k,l} = \frac{1}{\sqrt{p}}\mathrm{e}^{-j(l\omega+2\pi kl/p)},$$

where $k\in[0,p-1]$ and $l\in[0,p-1]$. The estimator $\hat{S}_{xy}^{(iii)}\left(k;\mathrm{e}^{j\omega}\right)$ is asymptotically unbiased and consistent, as long as $\hat{S}_{\underline{x}_p\underline{y}_p}\left(\mathrm{e}^{jp\omega}\right)$ is obtained from some asymptotically unbiased and consistent spectrum estimator, e.g., Welch-Barlett estimator and Thomson's estimator (see e.g., [68, 89]). In general, (4.24) can be proved by generalizing Theorem 5.1 appeared later that involves one cyclo-stationary signal. Even so, let us see (4.24) through a case of $p=2$ so that the mechanism of $\hat{S}_{xy}^{(iii)}\left(k;\mathrm{e}^{j\omega}\right)$ is explicitly revealed. That is, $\hat{S}_{xy}^{(iii)}\left(k;\mathrm{e}^{j\omega}\right)$ is reconstructed from all spectral components of $\underline{x}_p(l)$ and $\underline{y}_p(l)$ with spectrum cancelation.

Example 4.3 Let $x(t)$ and $y(t)$ be jointly $(\mathrm{CS})_2$. The correlation of the blocked signals $\underline{x}_2(l)$ and $\underline{y}_2(l)$ is associated with the time-varying correlation of $x(t)$ and $y(t)$:

$$\begin{aligned} r_{\underline{x}_2\underline{y}_2}(\tau) &= E\left\{\underline{x}_2(l+\tau)\,\underline{y}_2^*(l)\right\} \\ &= E\left\{\begin{bmatrix} x(2l+2\tau) \\ x(2l+2\tau+1)\end{bmatrix}\begin{bmatrix} y^*(2l) & y^*(2l+1)\end{bmatrix}\right\} \\ &= \begin{bmatrix} r_{xy}(0;2\tau) & r_{xy}(1;2\tau-1) \\ r_{xy}(0;2\tau+1) & r_{xy}(1;2\tau)\end{bmatrix}. \end{aligned}$$

Here the last equality utilizes the periodicity in (4.10). Then, the spectrum of $\underline{x}_2$ and $\underline{y}_2$ is

$$\begin{aligned} S_{\underline{x}_2\underline{y}_2}\left(\mathrm{e}^{j2\omega}\right) &= \sum_{\tau=-\infty}^{\infty} r_{\underline{x}_2\underline{y}_2}(\tau)\,\mathrm{e}^{-j2\omega\tau} \\ &= \sum_{n=-\infty}^{\infty}\frac{1}{2}\sum_{m=0}^{1}\mathrm{e}^{-j\pi mn}\begin{bmatrix} r_{xy}(0;n) & r_{xy}(1;n-1) \\ r_{xy}(0;n+1) & r_{xy}(1;n)\end{bmatrix}\mathrm{e}^{-j\omega n} \\ &= \frac{1}{2}\sum_{m=0}^{1}\sum_{n=-\infty}^{\infty}\begin{bmatrix} \sum_{k=0}^{1}R_{xy}(k;n) & \sum_{k=0}^{1}R_{xy}(k;n-1)\,\mathrm{e}^{j\pi k} \\ \sum_{k=0}^{1}R_{xy}(k;n+1) & \sum_{k=0}^{1}R_{xy}(k;n)\,\mathrm{e}^{j\pi k}\end{bmatrix}\mathrm{e}^{-j(\omega+\pi m)n} \\ &= \frac{1}{2}\sum_{m,k=0}^{1}S_{xy}(k;\omega+\pi m)\begin{bmatrix} 1 & \mathrm{e}^{-j(\omega+\pi(m-k))} \\ \mathrm{e}^{j(\omega+\pi m)} & \mathrm{e}^{j\pi k}\end{bmatrix}, \end{aligned}$$

[3]Note that ω is the normalized frequency and the sampling periods of $\underline{x}_p(l)$ and $\underline{y}_p(l)$ are p times those of $x(t)$ and $y(t)$ (see Section 2.3).

where the second equality performs a variable change $n := 2\tau$ based on the identity in (4.8). Thus,

$$\begin{aligned}
& U_2\left(\mathrm{e}^{j\omega}\right) \cdot S_{\underline{x}_2 \underline{y}_2}\left(\mathrm{e}^{j2\omega}\right) \cdot U_2^*\left(\mathrm{e}^{j\omega}\right) \\
= & \frac{1}{4}\begin{bmatrix} 1 & \mathrm{e}^{-j\omega} \\ 1 & -\mathrm{e}^{-j\omega} \end{bmatrix} \sum_{m,k=0}^{1} S_{xy}\left(k; \omega + \pi m\right) \begin{bmatrix} 1 & \mathrm{e}^{-j(\omega+\pi(m-k))} \\ \mathrm{e}^{j(\omega+\pi m)} & \mathrm{e}^{j\pi k} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \mathrm{e}^{j\omega} & -\mathrm{e}^{j\omega} \end{bmatrix} \\
= & \frac{1}{2}\begin{bmatrix} 1 & \mathrm{e}^{-j\omega} \\ 1 & -\mathrm{e}^{-j\omega} \end{bmatrix} \begin{bmatrix} S_{xy}(0;\omega) + S_{xy}(1;\omega+\pi) & S_{xy}(1;\omega) + S_{xy}(0;\omega+\pi) \\ \left(S_{xy}(0;\omega) - S_{xy}(1;\omega+\pi)\right)\mathrm{e}^{j\omega} & \left(S_{xy}(1;\omega) - S_{xy}(0;\omega+\pi)\right)\mathrm{e}^{j\omega} \end{bmatrix} \\
= & \begin{bmatrix} S_{xy}(0;\omega) & S_{xy}(1;\omega) \\ S_{xy}(1;\omega+\pi) & S_{xy}(0;\omega+\pi) \end{bmatrix},
\end{aligned}$$

whose first row consists of the cyclic spectra to be estimated. □

The next example validates performance of the smoothed cyclic periodogram and the blocking-based estimator, while the cyclic correlogram is omitted because of its equivalence to the smoothed cyclic periodogram.

Example 4.4 An AR system (see Example 5.3.2 in [89])

$$G(q) = \frac{1}{1 - 2.7607q^{-1} + 3.8104q^{-2} - 2.6535q^{-3} + 0.9238q^{-4}}$$

is driven by a $(\mathrm{CS})_4$ signal (see Example 4.1),

$$x(t) = \left(1 + \cos\left(\frac{2\pi t}{4}\right)\right) e(t),$$

where $e(t)$ is $\mathrm{WN}(0,1)$. As $G(q)$ is LTI, its output $y(t)$ is $(\mathrm{CS})_4$; the cyclic spectrum of $y(t)$ is (see (5.15) appeared later)

$$S_{yy}\left(k; \mathrm{e}^{j\omega}\right) = G\left(\mathrm{e}^{j\omega}\right) S_{xx}\left(k; \mathrm{e}^{j\omega}\right) G^*\left(\mathrm{e}^{j(\omega - 2\pi k/4)}\right), \tag{4.25}$$

where $S_{xx}\left(k; \mathrm{e}^{j\omega}\right)$ can be shown as

$$S_{xx}\left(k; \mathrm{e}^{j\omega}\right) = \begin{cases} 1.5, & k = 0 \\ j, & k = 1 \\ -0.5, & k = 2 \\ -j, & k = 3 \end{cases}, \ \forall \omega.$$

Figure 4.2 presents the true cyclic spectrum $S_{yy}\left(k; \mathrm{e}^{j\omega}\right)$ and its two estimates $\hat{S}_{yy}^{(L)}\left(k; \mathrm{e}^{j\omega}\right)$ and $\hat{S}_{yy}^{(\mathrm{iii})}\left(k; \mathrm{e}^{j\omega}\right)$. Both estimators capture most parts of $S_{yy}\left(k; \mathrm{e}^{j\omega}\right)$, with some discrepancies at low-magnitude areas. The discrepancies between $S_{yy}\left(k; \mathrm{e}^{j\omega}\right)$ and $\hat{S}_{yy}^{(\mathrm{iii})}\left(k; \mathrm{e}^{j\omega}\right)$ are larger, possibly due to the incomplete spectrum cancelation inherent in $\hat{S}_{yy}^{(\mathrm{iii})}\left(k; \mathrm{e}^{j\omega}\right)$. □

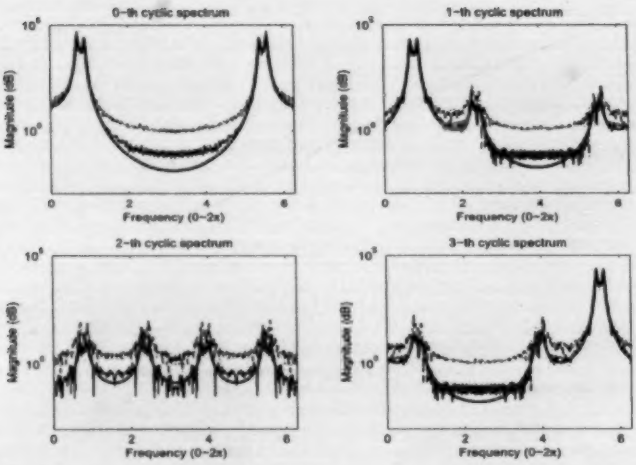

Figure 4.2: Estimate the cyclic spectrum ($N = 4096$): $S_{yy}\left(k;\mathrm{e}^{j\omega}\right)$ (smooth & solid), $\hat{S}_{yy}^{(\mathrm{i})}\left(k;\mathrm{e}^{j\omega}\right)$ (thick solid) and $\hat{S}_{yy}^{(\mathrm{iii})}\left(k;\mathrm{e}^{j\omega}\right)$ (dash).

4.5 Cyclo-Stationarity and Quasi-Stationarity

Estimators $\hat{m}_x^{(\mathrm{i})}(t)$, $\hat{r}_{xy}^{(\mathrm{i})}(t;\tau)$ and $\hat{S}_{xy}^{(\mathrm{iii})}\left(k;\mathrm{e}^{j\omega}\right)$ are based on the periodicities of cyclo-statistics and the blocked representations of cyclo-stationary signals; they can be conveniently implemented via the corresponding statistic estimators defined for stationary signals. This section shows that the other cyclo-statistic estimators can be similarly implemented via these stationary statistic estimators after a data transformation. The rationale of such an implementation shortcut is explored from the relationship between cyclo-stationarity and quasi-stationarity.

It is straightforward to see from (4.6) and (4.15) that $\hat{M}_x(k)$ and $\hat{R}_{xy}(k;\tau)$ can be implemented by applying stationary estimators of the mean and the correlation to $x(t)\mathrm{e}^{-j2\pi kt/p}$ and $(x(t),\ y(t)\mathrm{e}^{j2\pi kt/p})$, respectively. Similarly, the cyclic periodogram is equivalent to the stationary periodogram applying to $x(t)$ and $y(t)\mathrm{e}^{j2\pi kt/p}$, shown as follows. By defining $z(t) := y(t)\,\mathrm{e}^{j2\pi kt/p}$, (4.20) can be re-written as

$$\begin{aligned}
\hat{S}_{xy}^{(\mathrm{i})}\left(k;\mathrm{e}^{j\omega}\right) &= \frac{1}{N}X_N\left(\mathrm{e}^{j\omega}\right)Y_N^*\left(\mathrm{e}^{j(\omega-2\pi k/p)}\right) \\
&= \frac{1}{N}X_N\left(\mathrm{e}^{j\omega}\right)\left(\sum_{t=0}^{N-1}y(t)\,\mathrm{e}^{-j(\omega-2\pi k/p)t}\right)^* \\
&= \frac{1}{N}X_N\left(\mathrm{e}^{j\omega}\right)\left[\sum_{t=0}^{N-1}\left(y(t)\,\mathrm{e}^{j2\pi kt/p}\right)\mathrm{e}^{-j\omega t}\right]^*
\end{aligned}$$

$$= \frac{1}{N} X_N\left(e^{j\omega}\right) z_N^*\left(e^{j\omega}\right). \tag{4.26}$$

The right-hand side of (4.26) is the same as the stationary periodogram of $x(t)$ and $z(t)$, denoted as $\hat{S}_{xz}\left(e^{j\omega}\right)$, i.e.,

$$\hat{S}_{xy}^{(i)}\left(k; e^{j\omega}\right) = \hat{S}_{x(t),y(t)e^{j2\pi kt/p}}\left(e^{j\omega}\right). \tag{4.27}$$

The significance of (4.27) is not confined to a reliable and simple implementation of $\hat{S}_{xy}^{(i)}\left(k; e^{j\omega}\right)$ or $\hat{S}_{xy}^{(L)}\left(k; e^{j\omega}\right)$; more importantly, it implies that $S_{xy}\left(k; e^{j\omega}\right)$ can be estimated by all consistent spectrum estimators for stationary signals, e.g., Welch-Barlett's estimator and Thomson's estimator (see e.g., [68, 89]), not just $\hat{S}_{xy}^{(L)}\left(k; e^{j\omega}\right)$.

A question arises naturally: how to theoretically justify the shortcut of applying stationary statistic estimators to cyclo-stationary signals? The rationale lies on a fact that cyclo-stationary signals are quasi-stationary. Quasi-stationarity is defined in the same way as stationarity except that the expectation $E\{\cdot\}$ is replaced by [85]

$$\bar{E}\{\cdot\} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} E\{\cdot\}.$$

Quasi-stationarity has been shown to be a suitable unifying framework for signals appearing in the practice. It is an extension of stationarity and encloses more types of signals, e.g., stationary signals and deterministic periodic signals; meanwhile, quasi-stationary signals inherit many properties from stationary signals [85]. Specifically, the similarity between $E\{\cdot\}$ and $\bar{E}\{\cdot\}$ implies that quasi-stationary signals share the same statistic estimators with stationary signals. The following two equations show that cyclo-stationarity and quasi-stationarity are connected:

$$\begin{aligned}
\bar{m}_{x(t)e^{-j2\pi kt/p}} &= \bar{E}\left\{x(t)\, e^{-j2\pi kt/p}\right\} \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} m_x(t)\, e^{-j2\pi kt/p} \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \sum_{n=0}^{p-1} M_x(n)\, e^{j2\pi(n-k)t/p} \\
&= M_x(k)
\end{aligned} \tag{4.28}$$

and

$$\begin{aligned}
\bar{r}_{x(t),\, y(t)e^{j2\pi kt/p}}(\tau) &= \bar{E}\left\{x(t+\tau)\, y^*(t)\, e^{-j2\pi kt/p}\right\} \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} r_{xy}(t;\tau)\, e^{-j2\pi kt/p}
\end{aligned}$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \sum_{n=0}^{p-1} R_{xy}(n;\tau)\, e^{j2\pi(\alpha-k)n/p}$$
$$= R_{xy}(k;\tau). \tag{4.29}$$

Here the identity in (4.8) is used to achieve the last equalities of (4.28) and (4.29). Since $\hat{m}_{x(t)e^{j2\pi\alpha t/p}}$ and $\hat{r}_{x(t),y(t)e^{j2\pi\alpha t/p}}(\tau)$ are time independent, $x(t)$ and $y(t)e^{j2\pi\alpha t/p}$ are jointly quasi-stationary. Therefore, (4.28), (4.29) and

$$\hat{S}_{x(t),y(t)e^{j2\pi\alpha t/p}}\left(e^{j\omega}\right) = S_{xy}\left(k;e^{j\omega}\right),$$

which is implied by (4.29), are the theoretical counterparts of (4.6), (4.15) and (4.27), respectively.

Remark: Gardner [56] had an observation similar to (4.27),

$$S_{xx}\left(k;e^{j\omega}\right) = S_{x(t)e^{-j\pi\alpha t/p},x(t)e^{j\pi\alpha t/p}}\left(e^{j\omega}\right), \tag{4.30}$$

which was realized via the subband representation[4] of $x(t)$. Note that the rationale of (4.30) was not studied in [56].

4.6 Conclusion

This chapter studies the cyclo-statistic estimators, including the first class of estimators, $\hat{m}_x^{(t)}(t)$ in (4.4), $\hat{r}_{xy}^{(t)}(t;\tau)$ in (4.13), $\hat{S}_{xy}^{(mA)}\left(k;e^{j\omega}\right)$ in (4.23) and $\hat{S}_{xy}^{(cm)}\left(k;e^{j\omega}\right)$ in (4.26), and the second class of estimators, $\hat{m}_x^{(n)}(t)$ in (4.5), $\hat{r}_{xy}^{(n)}(t;\tau)$ in (4.14), $\hat{R}_{xy}(k;\tau)$ in (4.15), $\hat{S}_{xy}^{(t)}\left(k;e^{j\omega}\right)$ in (4.21) and $\hat{S}_{xy}^{(nB)}\left(k;e^{j\omega}\right)$ in (4.22). The second class of estimators has an implementation shortcut, e.g., (4.27) for the cyclic periodogram. The rationale of the shortcut is explored from the fact that cyclo-stationary signals are quasi-stationary (see (4.28) and (4.29)).

[4] The subband representation of a discrete-time cyclo-stationary signal was formally proposed in [113, 61].

Chapter 5

Cyclo-Spectral Theory

This chapter[1] studies two problems in the spectral theory of discrete-time cyclo-stationary signals: the cyclospectrum representation and the cyclospectrum transformation by linear multirate systems. Four types of cyclospectra are presented and their interrelationships are explored. In the literature, the problem of cyclospectrum transformation by linear systems was investigated only for some specific configurations and was usually developed with inordinate complexities due to lack of a systematic approach. A general multirate system that encompasses most common systems — linear time-invariant systems and linear periodically time-varying systems — is proposed as the unifying framework; more importantly, it also includes many configurations that have not been investigated before, e.g., fractional sample-rate changers with cyclo-stationary inputs. The blocking technique provides a systematic solution as it associates a multirate system with an equivalent linear time-invariant system, and cyclo-stationary signals with stationary signals; thus, the original problem is elegantly converted into a relatively simple one, which is solved in the form of matrix multiplication.

5.1 Introduction

The spectral theory of cyclo-stationary signals has applications in different areas, e.g., blind channel identification and equalization by fractional sampling received signals [143, 144], filter bank optimization by minimizing averaged variances of reconstruction errors [112, 99], system identification by introducing cyclo-stationary external excitation [55, 60] and by fast sampling system outputs [132, 151].

The spectral theory of discrete-time cyclo-stationary signals mainly consists of two parts, namely, the cyclospectrum representation and the cyclospectrum transformation by linear systems. Here the cyclospectrum is the counterpart of the power spectrum defined for

[1]The chapter has been published in [152].

discrete-time stationary or strictly speaking wide-sense-stationary signals. The theory was first developed by Gladyshev [62]: a complex function that is currently referred to as the cyclic spectrum was defined as the spectrum of a p-periodically correlated[2] sequence; the spectral relationship between the original sequence and a higher dimensional sequence that is actually the blocked signal was discussed. Motivated by the sampling operation, the cyclic spectrum of discrete-time cyclo-stationary signals was defined but only a very limited study has been given in Gardner's books [54, 53]; as a complement, LTI and linear periodically time-varying (LPTV) filtering of cyclo-stationary signals was discussed briefly in [56]. Using the Gardner's notation (e.g., that in [54]), Ohno and Sakai [99] derived the output cyclic spectrum of a filter bank (an LPTV system) mostly from definitions and used it in the optimal filter bank design. To avoid the cumbersome derivation in [99], Sakai and Ohno [112] studied the cyclic spectrum relationships among the original, the modulated, and the blocked signals, and obtained the same expression of the cyclic spectrum in [99] via these relationships. In an excellent overview [61], Giannakis presented some results in terms of the cyclic spectrum on the LPTV filtering, fractional sampling and multirate processing. Besides the cyclic spectrum, there are some other cyclospectra, namely, the time frequency representation (TFR), the bispectrum and the 2-D spectrum. After giving an observation that the cyclic spectrum is not "very illustrative" (a character actually caused by derivation without a systematic approach), Lall et al. [79] analyzed the output of a filter bank in terms of the TFR. Akkarakaran and Vaidyanathan [2] used the bispectrum as a tool to generalize most results in [113] (studying effects of multirate blocks on scalar cyclo-stationary signals) into the vector case; they also gave the bispectrum of the output of a single-input and single-output (SISO) LPTV system and found the conditions under which a SISO LPTV system would produce stationary outputs for all stationary inputs. The 2-D spectrum, indeed a coordinate transform of the bispectrum, was proposed in the context of periodic random processes in [136, 138, 137] where it was related to the cyclic spectrum and the TFR.

These four types of cyclospectra should have some interrelationships, since they all describe second-order statistical properties of cyclo-stationary signals. The first contribution of this chapter, which is also of some tutorial value, is to summarize these cyclospectra and find their interrelationships. As shown later, they are indeed related to each other and mutually convertible, even though each has its own features and may be superior to others in one specific context or other.

The kernel problem of the spectral theory is the cyclospectrum transformation by linear

[2] "Periodically correlated" is a synonym of "cyclo-stationary" mainly used in the mathematical field [41].

systems, i.e., to represent the cyclospectrum of the system output in terms of that of the input. For this problem, there exist two limitations in the above cited literature: first, only some specific configurations have been investigated, e.g., the input of an p-band filter bank has to be either stationary [99, 112] or cyclo-stationary with the same period as p [79]; second, most of existing results, e.g., the cyclic spectrum of the output of an LPTV system (Eq. (17.44) in [61]), are developed via definitions and hence their derivations are so overwhelming that generalization to more complex systems, e.g., multirate systems, becomes almost impossible unless a systematic approach is adopted (see the proof of (5.15) in Example 5.4). The second contribution of this chapter is to remove these limitations: the problem of the cyclospectrum transformation is attacked in the framework of multirate systems using the blocking technique.

A discrete-time linear system can always be represented by a Green's function $g(k,l)$ as

$$y(k) = \sum_{l=-\infty}^{\infty} g(k,l)\, x(l), \quad \forall k, \tag{5.1}$$

where $k, l \in \mathbb{Z}$ [35]. A linear SISO multirate system[3] G has the so-called (m,n)–shift invariance property (m and n are integers) if shifting the input by n samples results in shifting the output by m samples [29]. In terms of the Green's function, (m,n)–shift invariance is characterized by

$$g(k+m, l+n) = g(k,l), \quad \forall k, l. \tag{5.2}$$

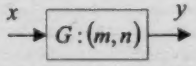

Figure 5.1: A linear SISO multirate system

Figure 5.1 depicts such a linear SISO multirate system G, in which the notation "$G : (m,n)$" denotes that G is (m,n)-shift invariant. Such a multirate system G covers many familiar systems as special cases, e.g., the LTI system ($m = n = 1$), the LPTV system ($m = n$), and the cascade of upsampler, LTI system and downsampler (m and n are coprime) depicted in Figure 5.2.

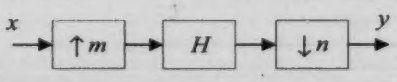

Figure 5.2: A cascade of upsampler ($\uparrow m$), LTI system (H) and downsampler ($\downarrow n$)

[3]SISO multirate systems are also called dual-rate systems [29].

Blocking has been shown to be a powerful technique in dealing with cyclo-stationary signals and multirate systems: blocking the cyclo-stationary signal can result in a higher dimensional stationary signal (see Section 2.3); by the blocking technique, one can associate the multirate system with an equivalent multi-input and multi-output LTI system [96, 78].

Therefore, our main idea is to block multirate systems and cyclo-stationary signals properly and convert the original problem into one involving LTI systems and stationary signals only that can be readily solved using some well-known results. More specifically, the kernel problem is separated into the following two sub-questions:

- Given a linear SISO multirate system $G : (m, n)$ in Figure 5.1 and the input x is $(\mathrm{CS})_p$, is the output y stationary or cyclo-stationary? If y is cyclo-stationary, what is its period?
- What is the cyclospectrum transformation in Figure 5.1, i.e., how to represent the cyclospectrum of y in terms of that of x?

The rest of the chapter is organized as follows. Section 5.2 summarizes the different cyclospectra and explores their interrelationships. Section 5.3 studies the effects of the blocking operation on statistical properties of cyclo-stationary signals. Section 5.4 answers the two sub-questions and presents some examples as illustration. Finally, Section 5.5 provides concluding remarks.

5.2 Cyclospectrum

We first review four types of cyclospectra and show that these cyclospectra are related to each other. After presenting the cyclospectrum transformation by LTI systems, we choose the cyclic spectrum as the representation of the cyclospectrum in the rest sections.

There are mainly four types of cyclospectra, namely, the cyclic spectrum, the time frequency representation, the bispectrum and the 2-D spectrum.

Cyclic Spectrum: To parallel with other cyclospectra, we introduce the cyclic spectrum that appeared in previous chapters again with slightly different notation. Denote the time-varying correlation by $R_{xx}(t+\tau, t) := E\{x(t+\tau)x^*(t)\}$. Eqs. (2.6) and (2.7) say that $R_{xx}(t+\tau, t)$ is a periodic sequence of t with period p for a fixed τ. So $R_{xx}(t+\tau, t)$ has the following discrete Fourier expansion

$$R_{xx}(t+\tau, t) = \sum_{k=0}^{p-1} C_{xx}^{(k)}(\tau)\, e^{j2\pi kt/p}, \tag{5.3}$$

where the discrete Fourier series coefficient $C_{xx}^{(k)}(\tau)$ is

$$C_{xx}^{(k)}(\tau) = \frac{1}{p}\sum_{t=0}^{p-1} R_{xx}(t+\tau, t)\, \mathrm{e}^{-j2\pi kt/p}. \tag{5.4}$$

The DTFT of $C_{xx}^{(k)}(\tau)$ is defined as the cyclic spectrum of x [62, 53],

$$S_{xx}^{(k)}\left(\mathrm{e}^{j\omega}\right)_{CPS} = \sum_{\tau=-\infty}^{\infty} C_{xx}^{(k)}(\tau)\, \mathrm{e}^{-j\omega\tau}, \tag{5.5}$$

where the subscript $(_{CPS})$ stands for "cyclic power spectrum". In the sequel, the symbol $\left(\mathrm{e}^{j\omega}\right)$ is often shorten as (ω) for simple notation. It follows from (2.7) and (5.4) that $S_{xx}^{(k)}(\omega)_{CPS}$ is periodic in k with period p,

$$S_{xx}^{(k+lp)}(\omega)_{CPS} = S_{xx}^{(k)}(\omega)_{CPS}, \ \forall k, l \in \mathbb{Z}. \tag{5.6}$$

The set $\left\{S_{xx}^{(0)}(\omega)_{CPS}, S_{xx}^{(1)}(\omega)_{CPS}, \cdots, S_{xx}^{(p-1)}(\omega)_{CPS}\right\}$ thus forms a full description of the cyclospectrum of a $(\mathrm{CS})_p$ signal x. The cyclic spectrum can be considered as a generalization of the power spectrum in (2.3); that is, if x is stationary, $S_{xx}^{(0)}(\omega)_{CPS} = S_{xx}(\omega)$ and $S_{xx}^{(k)}(\omega)_{CPS} = 0$ for $k = 1, 2, \cdots, p-1$.

Time Frequency Representation: $R_{xx}(t+\tau, t)$ can be considered also as a sequence of τ for a fixed t. The time frequency representation is defined as the DTFT of $R_{xx}(t+\tau, t)$ taking τ as the changing variable [79],

$$S_{xx}^{(t)}(\omega)_{TFR} = \sum_{\tau=-\infty}^{\infty} R_{xx}(t+\tau, t)\mathrm{e}^{-j\omega\tau}. \tag{5.7}$$

It follows from (2.7) that $S_{xx}^{(t)}(\omega)_{TFR}$ is periodic in t with period p,

$$S_{xx}^{(t+lp)}(\omega)_{TFR} = S_{xx}^{(t)}(\omega)_{TFR}, \ \forall t, l \in \mathbb{Z}.$$

If $p = 1$, i.e., x is stationary, $S_{xx}^{(t)}(\omega)_{TFR} = S_{xx}(\omega)$, for all t. The TFR is a broad concept that characterizes non-stationary signals over a jointly time-frequency domain [21, 69]. It is also known as Rihaczek spectrum [108] or the time-varying spectrum [114].

Bispectrum: The bispectrum $S_{xx}(\omega_1, \omega_2)$ is defined as the two-dimensional DTFT of the autocorrelation $R_{xx}(t_1, t_2)$ [2, 100]

$$S_{xx}(\omega_1, \omega_2) = \frac{1}{2\pi}\sum_{t_1=-\infty}^{\infty}\sum_{t_2=-\infty}^{\infty} R_{xx}(t_1, t_2)\, \mathrm{e}^{-j\omega_1 t_1}\mathrm{e}^{j\omega_2 t_2}. \tag{5.8}$$

Like the TFR, the bispectrum is also a very general concept that can describe the second-order statistical property of non-stationary signals. Specifically, the bispectrum of a cyclostationary signal lies on some parallel lines in the ω_1-ω_2 plane [2],

$$\omega_2 - \omega_1 + \frac{2\pi k}{p} = 0, \tag{5.9}$$

where $k \in \mathbb{Z}$. It is shown later that the bispectrum component on the k-th line is exactly the k-th cyclic spectrum defined in (5.5). Note that the terminology "bispectrum" has a different meaning in the literature — the two-dimensional DTFT of the third-order moment [135].

2-D Spectrum: The 2-D spectrum is defined as the two-dimensional Fourier transform of $R_{xx}(t+\tau, t)$ [136, 138, 137],

$$S_{xx}^{2D}(\lambda, \omega) = \sum_{\tau=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} R_{xx}(t+\tau, t)\, \mathrm{e}^{-j\lambda t} \mathrm{e}^{-j\omega\tau}. \tag{5.10}$$

The 2-D spectrum and bispectrum are very similar (see (5.12) later); however, for a cyclostationary signal x, $S_{xx}^{2D}(\lambda, \omega)$ is continuous in ω and discrete in λ, but $S_{xx}(\omega_1, \omega_2)$ in (5.8) is continuous both in ω_1 and ω_2. The 2-D spectrum is also referred to as the dual-frequency spectrum [114] defined for non-stationary signals.

These cyclospectra are related to each other. First, it follows from (5.4), (5.5) and (5.7) that the cyclic spectrum and the TFR are a discrete Fourier transform pair [79]

$$S_{xx}^{(k)}(\omega)_{CPS} = \frac{1}{p} \sum_{t=0}^{p-1} S_{xx}^{(t)}(\omega)_{TFR}\, \mathrm{e}^{-j2\pi kt/p}. \tag{5.11}$$

Second, it follows easily from (5.8) and (5.10) that the bispectrum and the 2-D spectrum are a coordinate transform of each other with a scaling factor,

$$S_{xx}^{2D}(\lambda, \omega) = 2\pi S_{xx}(\omega, \omega - \lambda). \tag{5.12}$$

Third, the k-th cyclic spectrum is exactly the bispectrum component that lies on the k-th line described in (5.9), i.e.,

$$S_{xx}(\omega_1, \omega_2) = \sum_{k=0}^{p-1} S_{xx}^{(k)}(\omega_1)_{CPS}\, \delta\left(\omega_2 - \omega_1 + \frac{2\pi k}{p} + 2\pi l\right), \tag{5.13}$$

where l is an integer and $\delta(\cdot)$ denotes the Dirac delta function [61]. Certainly, the k-th cyclic spectrum is also related to the 2-D spectrum [138]

$$S_{xx}^{(k)}(\omega)_{CPS} = \frac{1}{2\pi} S_{xx}^{2D}\left(\frac{2\pi k}{p}, \omega\right).$$

Finally, the TFR and the 2-D spectrum are a DTFT pair [136],

$$S_{xx}^{2D}(\lambda, \omega) = \sum_{t=-\infty}^{\infty} S_{xx}^{(t)}(\omega)_{TFR}\, \mathrm{e}^{-j\lambda t}. \tag{5.14}$$

These interrelationships are shown in Figure 5.3.

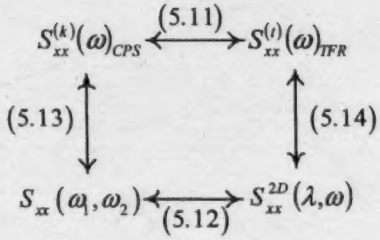

Figure 5.3: Interrelationships among the four cyclospectra

Since these cyclospectra are mutually convertible, it is sufficient to use only one of them to represent the cyclospectrum in the rest of the development. Before making a choice, we introduce the cyclospectrum transformation by LTI systems to further capture the characteristics of these cyclospectra.

Let a $(\mathrm{CS})_p$ signal x be the input of a discrete-time LTI system G with transfer function $\hat{G}(z)$. As LTI systems preserve the cyclo-stationarity [113, 79, 2], the output y is $(\mathrm{CS})_p$. The cyclospectrum of y is associated with that of x as follows. In terms of the cyclic spectrum,

$$S_{yy}^{(k)}(\omega)_{CPS} = \hat{G}\left(e^{j\omega}\right) S_{xx}^{(k)}(\omega)_{CPS} \hat{G}^*\left(e^{j(\omega-\omega_0 k)}\right), \tag{5.15}$$

where $\omega_0 = 2\pi/p$. Similar observations to (5.15) were noticed in [56, 144] without proofs, which are provided in Example 5.4. The TFR and bispectrum of y were given in [79] and [2], respectively,

$$S_{yy}^{(t)}(\omega)_{TFR} = \hat{G}\left(e^{j\omega}\right) \sum_{l=-\infty}^{\infty} S_{xx}^{(t-l)}(\omega)_{TFR}\, g^*(l)\, e^{j\omega l}, \tag{5.16}$$

$$S_{yy}(\omega_1,\omega_2) = \hat{G}\left(e^{j\omega_1}\right) S_{xx}(\omega_1,\omega_2) \hat{G}^*\left(e^{j\omega_2}\right). \tag{5.17}$$

Here $g(\cdot)$ is the impulse response of G. Eqs. (5.12) and (5.17) give the 2-D spectrum of y,

$$S_{yy}^{2D}(\lambda,\omega) = \hat{G}\left(e^{j\omega}\right) S_{xx}^{2D}(\lambda,\omega) \hat{G}^*\left(e^{j(\omega-\lambda)}\right). \tag{5.18}$$

The TFR is introduced because "the representation of the cyclo-stationary processes, in terms of cyclic spectral density[4], which, although a means of characterization, is not very illustrative, particularly in the context of filter bank analysis" [79]; however, (5.16) reveals that the TFR is not compact. The bispectrum is originally defined for non-stationary signals so that it has been cautioned to be unwieldy in mathematics [103] or too general and

[4]The cyclic spectral density is synonymous to the cyclic spectrum.

inefficient for indiscriminate use [2]. The 2-D spectrum is simply a coordinate transformation of the bispectrum, and thus they share the same problems. The cyclic spectrum is defined in the way of incorporating the spectral information along with periodicity and thus it displays directly the fundamental characteristic of cyclo-stationary signals; even though (5.15) is not as compact as (5.17) and (5.18), the cyclic spectrum is very convenient once all the k-th cyclic spectra are enclosed in the so-called cyclic spectrum matrix [62, 112]. Therefore, we will use the cyclic spectrum in the sequel and the subscript ($_{CPS}$) is dropped without confusion.

5.3 Blocking Operator

In Section 5.1 we have briefly discussed the idea of blocking multirate systems and cyclo-stationary signals properly to form a relatively simple problem. Based on Section 2.3, this section studies further the effects of the blocking operator on multirate systems and cyclo-stationary signals.

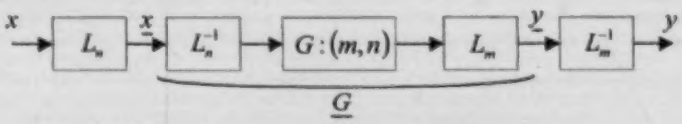

Figure 5.4: Blocking a linear SISO multirate system

One of the advantages of the blocking operator is that it can associate multirate systems that are essentially time varying with some equivalent LTI systems to which many existing LTI techniques can be applied. For the multirate system in Figure 5.1, blocking the input x and the output y by L_n and L_m respectively yields a blocked system $\underline{G} := L_m G L_n^{-1}$, which has n inputs and m outputs. The blocking procedure is depicted in Figure 5.4. As G is (m,n)-shift invariant (see (5.2)), $\underline{G}$ is LTI [96] and has an $m \times n$ transfer matrix [29]

$$\hat{\underline{G}}(z) = \begin{bmatrix} \hat{G}_{00}(z) & \hat{G}_{01}(z) & \cdots & \hat{G}_{0,n-1}(z) \\ \hat{G}_{10}(z) & \hat{G}_{11}(z) & \cdots & \hat{G}_{1,n-1}(z) \\ \vdots & \vdots & & \vdots \\ \hat{G}_{m-1,0}(z) & \hat{G}_{m-1,1}(z) & \cdots & \hat{G}_{m-1,n-1}(z) \end{bmatrix}, \tag{5.19}$$

whose entries relate to the Green's function of G (see (5.1)) as

$$\hat{G}_{k,l}(z) = \sum_{t=-\infty}^{\infty} g(k+mt, l)\, z^{-t}. \tag{5.20}$$

Gladyshev [62] first proposed the relationship between a cyclic spectrum matrix of the original signal and the power spectrum of the blocked signal, which was proved by Sakai

and Ohno [112] for scalar signals. Akkarakaran and Vaidyanathan [2] also noticed this kind of relationship between bispectra. Here we offer a theorem describing this relationship in terms of the cyclic spectrum for vector signals.

Theorem 5.1 *The cyclic spectrum matrix $\Phi_{xx}(\omega)$ of a $(CS)_p$ q-dimensional vector signal x is connected with the power spectrum $S_{\underline{x}\underline{x}}(\omega)$ of its p-fold blocked version $\underline{x}$ as*

$$\Phi_{xx}(\omega) = U_{p|q}(\omega) S_{\underline{x}\underline{x}}(p\omega) U_{p|q}^{*}(\omega), \tag{5.21}$$

where $\Phi_{xx}(\omega)$ is an $qp \times qp$ matrix whose kl-th $q \times q$ block component is determined by the cyclic spectrum of x,

$$[\Phi_{xx}(\omega)]_{kl} = S_{xx}^{(k-l)}(\omega + k\omega_0), \tag{5.22}$$

and $U_{p|q}(\omega)$ is an $qp \times qp$ unitary matrix whose kl-th $q \times q$ block component is

$$\left[U_{p|q}(\omega)\right]_{kl} = \frac{1}{\sqrt{p}} e^{-j(kl\omega_0 + l\omega)} I_q. \tag{5.23}$$

Here $k = 0, 1, \cdots, p-1$, $l = 0, 1, \cdots, p-1, \omega_0 = 2\pi/p$ and I_q is a $q \times q$ identity matrix.

Proof of Theorem 5.1: Since the blocked signal $\underline{x}_p$ is stationary (see Section 2.3), it has a power spectrum $S_{\underline{x}\underline{x}}(\omega)$. From (2.3) and (2.8), the kl-th component of $S_{\underline{x}\underline{x}}(p\omega)$ is

$$\begin{aligned}
[S_{\underline{x}\underline{x}}(p\omega)]_{kl} &= \sum_{\tau=-\infty}^{\infty} E\left[x_k(t+\tau) x_l^{*}(t)\right] e^{-jp\omega\tau} \\
&= \sum_{\tau=-\infty}^{\infty} E\left[x(pt+p\tau+k) x^{*}(pt+l)\right] e^{-jp\omega\tau}.
\end{aligned}$$

An identity

$$\frac{1}{p}\sum_{m=0}^{p-1} e^{-j\omega_0 mv} = \begin{cases} 1, & v/p \text{ is an integer,} \\ 0, & \text{otherwise,} \end{cases}$$

is used to change $p\tau$ by a new variable v,

$$\begin{aligned}
[S_{\underline{x}\underline{x}}(p\omega)]_{kl} &= \sum_{v=-\infty}^{\infty} \frac{1}{p} \sum_{m=0}^{p-1} e^{-j\omega_0 mv} R_{xx}(pt+v+k, pt+l) e^{-j\omega v} \\
&= \frac{1}{p} \sum_{\tau=-\infty}^{\infty} \sum_{m=0}^{p-1} e^{-j\omega_0 m(\tau+l-k)} R_{xx}(pt+\tau+l, pt+l) e^{-j\omega(\tau+l-k)} \\
&= \frac{1}{p} \sum_{\tau=-\infty}^{\infty} \sum_{m=0}^{p-1} e^{-j\omega_0 m(\tau+l-k)} R_{xx}(l+\tau, l) e^{-j\omega(\tau+l-k)},
\end{aligned}$$

where the second equality is reached by replacing v with a variable $\tau := v + k - l$ and the last equality follows from the cyclo-stationarity property described in (2.7). From (5.3) and (5.5),

$$\begin{aligned}
[S_{\underline{x}\underline{x}}(p\omega)]_{kl} &= \frac{1}{p}\sum_{\tau=-\infty}^{\infty}\sum_{m=0}^{p-1} \mathrm{e}^{-j\omega_0 m(\tau+l-k)} \sum_{r=0}^{p-1} C_{xx}^{(r)}(\tau)\, \mathrm{e}^{j\omega_0 rl} \mathrm{e}^{-j\omega(\tau+l-k)} \\
&= \frac{1}{p}\sum_{m=0}^{p-1}\sum_{r=0}^{p-1} \mathrm{e}^{j(\omega_0 m+\omega)k} \mathrm{e}^{-j\omega_0(m-r)l} \mathrm{e}^{-j\omega l} S_{xx}^{(r)}(\omega+\omega_0 m). \qquad (5.24)
\end{aligned}$$

Next, we show that (5.24) is equivalent to

$$[S_{\underline{x}\underline{x}}(p\omega)]_{kl} = \frac{1}{p}\sum_{m=0}^{p-1}\sum_{n=0}^{p-1} \mathrm{e}^{j(\omega_0 m+\omega)k} \mathrm{e}^{-j(\omega_0 n+\omega)l} S_{xx}^{(m-n)}(\omega+\omega_0 m). \qquad (5.25)$$

Comparing (5.24) and (5.25), their difference for a certain $m \in [0, p-1]$ is

$$\begin{aligned}
&\sum_{r=0}^{p-1} \mathrm{e}^{-j\omega_0(m-r)l} S_{xx}^{(r)}(\omega+\omega_0 m) - \sum_{n=0}^{p-1} \mathrm{e}^{-j\omega_0 nl} S_{xx}^{(m-n)}(\omega+\omega_0 m) \\
= &\sum_{r=0}^{p-1} \mathrm{e}^{-j\omega_0(m-r)l} S_{xx}^{(r)}(\omega+\omega_0 m) \\
&- \left[\sum_{n=0}^{m} \mathrm{e}^{-j\omega_0 nl} S_{xx}^{(m-n)}(\omega+\omega_0 m) + \sum_{n=m+1}^{p-1} \mathrm{e}^{j\omega_0(p-n)l} S_{xx}^{(p+m-n)}(\omega+\omega_0 m)\right] \\
= &\sum_{r=0}^{p-1} \mathrm{e}^{-j\omega_0(m-r)l} S_{xx}^{(r)}(\omega+\omega_0 m) \\
&- \left[\sum_{r=0}^{m} \mathrm{e}^{-j\omega_0(m-r)l} S_{xx}^{(r)}(\omega+\omega_0 m) + \sum_{r=m+1}^{p-1} \mathrm{e}^{-j\omega_0(m-r)l} S_{xx}^{(r)}(\omega+\omega_0 m)\right] \\
= &\ 0,
\end{aligned}$$

where the second equality uses the periodicity of the cyclic spectrum in (5.6) and the third equality is obtained by changing variables $r := m - n$ and $r := p + m - n$ in the last two summing terms, respectively. Finally, (5.21) is obtained from (5.25). □

Remark: There are two important differences between Theorem 5.1 and its counterpart in [112]: first, the proof in [112] takes a modulation representation of cyclo-stationary signals as an intermediate step, whereas we attack the problem more directly and thus the proof is much simpler; second, the result in [112] holds under a condition

$$|\omega| < \omega_0/2, \qquad (5.26)$$

which is actually introduced by the modulation representation, whereas our proof shows that (5.26) is superfluous. Removing the limiting condition (5.26) is extremely important,

because the valid range of using the blocking technique will become too small to be meaningful, if (5.26) has to be satisfied.

5.4 Cyclo-Stationary Signals in Multirate Systems

We are ready to attack the two sub-questions proposed in Section 5.1 using the blocking technique. The first sub-question is answered in Theorem 5.2; the second is solved in the form of matrix multiplication. Both are followed by some specific configurations as illustration.

5.4.1 Cyclo-Stationarity of the Output

There are basically three approaches to answer the first sub-question: (i) to explicitly write out some statistics of the system output, e.g., the autocorrelation or the cyclic spectrum, as what was done in [2, 79]; (ii) to reduce the multirate system into simpler building blocks such as: upsamplers, LTI systems and downsamplers and study cyclo-stationary properties of each block, as in [2, 113]; (iii) to use the blocking technique. The last approach is the simplest for most systems and hence is adopted here.

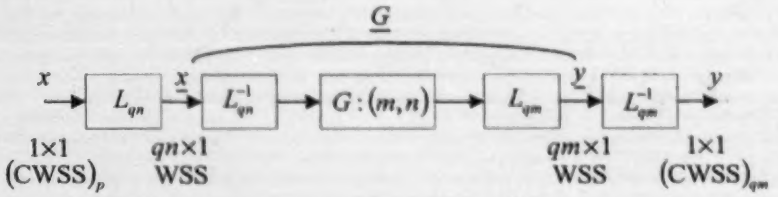

Figure 5.5: A blocked SISO multirate system

Theorem 5.2 *Given a linear SISO multirate system $G:(m,n)$ in Figure 5.1 and an $(CS)_p$ input x, the output y is $(CS)_{pm/\gcd(p,n)}$.*

Proof of Theorem 5.2: The use of the blocking operator needs to comply with two principles: the blocked-system needs to be LTI and the blocked signal is stationary. Thus, the fold of the blocking operator at the input side must be an integer multiple of n as $L_m G L_n^{-1}$ is LTI (see Figure 5.4) and at the same time be an integer multiple of p. Blocking x by L_{qn} and y by L_{qm} as depicted in Figure 5.5 will satisfy the two principles, where

$$q = p/\gcd(p,n). \tag{5.27}$$

In Figure 5.5, the blocked system $\underline{G} := L_{qm} G L_{qn}^{-1}$ is LTI and the blocked input $\underline{x}$ is stationary. The blocked output $\underline{y}$ is stationary, which implies that y is $(\mathrm{CS})_{pm/\gcd(p,n)}$. □

Example 5.1 The multirate system G is indeed an LTI system if $m = n = 1$, under which Theorem 5.2 says that if the input is $(\mathrm{CS})_p$, then the output is $(\mathrm{CS})_p$ too. In other words, LTI systems preserve the cyclo-stationarity, which is consistent with the conclusions in [113, 79, 2, 61]. □

Example 5.2 If $m = n \neq 1$, the multirate system $G : (m, n)$ reduces to an LPTV system that appears frequently in signal processing and control, such as multirate filter banks [147] and LPTV controllers [78, 50]. More specifically, if the input is stationary or $(\mathrm{CS})_p$, Theorem 5.2 gives that the output of an LPTV system with period p (i.e., $m = n = p$) is $(\mathrm{CS})_p$. This conclusion is consistent with those in [2, 113, 61]. □

Examples 5.1 and 5.2 are both with $m = n$. For $m \neq n$, the multirate system is also named the fractional sample-rate changer, which has two configurations as follows. First, if m and n are coprime, the multirate system G is equivalent to the cascade of upsampler, LTI system and downsampler, depicted in Figure 5.2, which has been studied extensively [147, 87, 113, 48, 118]. Second, if m and n have some nontrivial common factor, G is not equivalent to the cascade system in Figure 5.2 [118, 29]; it stands for a more general building block that finds applications in the nonuniform filter banks [29] and the multichannel nonuniform transmultiplexers [83].

Example 5.3 For the cascade system in Figure 5.2 (m and n are coprime), if the input x is stationary ($p = 1$), Theorem 5.2 says that the output y is $(\mathrm{CS})_m$, which is consistent with that in [113] obtained by analyzing the cascade in Figure 5.2. As a comparison, the other two approaches mentioned at the beginning of this subsection are explored for the same configuration. Clearly the first approach has difficulties as the explicit statistical expression of the system output has not been given in the literature. The second approach proceeds as follows. The upsampler and downsampler have the properties: if the input of a k-fold upsampler is $(\mathrm{CS})_p$, the output will be $(\mathrm{CS})_{kp}$ [79]; if the input of a k-fold downsampler is $(\mathrm{CS})_p$, the output is $(\mathrm{CS})_{p/\gcd(p,k)}$ [113]. Applying the two properties to the cascade system in Figure 5.2 gives the same result. □

Remark: Theorem 5.2 can be verified by estimating the period of the output via some numerical methods, e.g., Hurd-Gerr's method [72], Martin-Kedem's method [90], Dandawaté-Giannakis's method [37] and the variability method (Chapter 3).

5.4.2 Cyclospectrum of the Output

We follow the same idea used in Section 5.4.1. Specifically, the multirate system and cyclostationary signals are blocked as that in Figure 5.5 (see the proof of Theorem 5.2). Blocking the $(\mathrm{CS})_p$ input x by L_{qn} for the q in (5.27) implies that $pr = qn$, where r is an integer. A general solution of the second sub-question consists of two cases.

Case 1: $r = 1$. From Theorem 5.1, we have (see Figure 5.5)

$$S_{\underline{x}\,\underline{x}}(p\omega) = U_{p|1}^{*}(\omega)\,\Phi_{xx}(\omega)\,U_{p|1}(\omega), \tag{5.28}$$

$$\Phi_{yy}(\omega) = U_{qm|1}(\omega)\,S_{\underline{y}\,\underline{y}}(qm\omega)\,U_{qm|1}^{*}(\omega), \tag{5.29}$$

where $q = p/n$. Since $\underline{x}$ is stationary and $\underline{G} := L_{qm} G L_{qn}^{-1}$ is LTI, (2.4) gives

$$S_{\underline{y}\,\underline{y}}(\omega) = \hat{\underline{G}}\left(\mathrm{e}^{j\omega}\right) S_{\underline{x}\,\underline{x}}(\omega)\,\hat{\underline{G}}^{*}\left(\mathrm{e}^{j\omega}\right), \tag{5.30}$$

where $\hat{\underline{G}}(z)$ is represented by the Green's function of G in (5.19). Therefore, the cyclic spectrum of y is associated with that of x via (5.28), (5.29) and (5.30),

$$\Phi_{yy}(p\omega) = U_{qm|1}(p\omega)\,\hat{\underline{G}}\left(\mathrm{e}^{jpqm\omega}\right) U_{p|1}^{*}(qm\omega)\,\Phi_{xx}(qm\omega)\,U_{p|1}(qm\omega)\,\hat{\underline{G}}^{*}\left(\mathrm{e}^{jpqm\omega}\right) U_{qm|1}^{*}(p\omega). \tag{5.31}$$

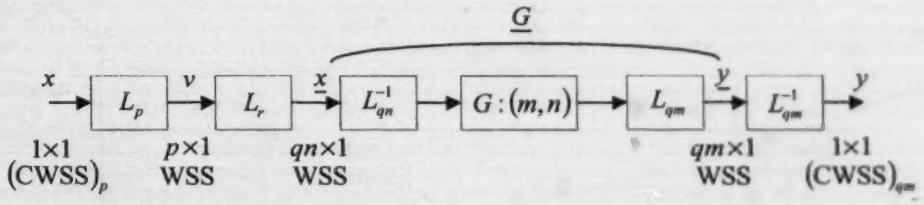

Figure 5.6: An equivalent blocked SISO multirate system: $pr = qn$

Case 2. $r > 1$*:* With $q = p/\gcd(p,n)$, (5.29) and (5.30) still hold. However, one more step is needed: L_{qn} is decomposed into a series of L_p and L_r, which makes Figure 5.5 equivalent to Figure 5.6. Theorem 5.1 gives

$$S_{vv}(p\omega) = U_{p|1}^{*}(\omega)\,\Phi_{xx}(\omega)\,U_{p|1}(\omega). \tag{5.32}$$

Since v is stationary, its cyclic spectrum matrix is block diagonal. Taking it as a special case of Theorem 5.1 results

$$S_{\underline{x}\,\underline{x}}(r\omega) = U_{r|p}^{*}(\omega)\cdot \mathrm{diag}\left(\begin{bmatrix} S_{vv}(\omega) \\ S_{vv}(\omega + 2\pi/r) \\ \vdots \\ S_{vv}(\omega + 2\pi(r-1)/r) \end{bmatrix}\right)\cdot U_{r|p}(\omega), \tag{5.33}$$

where $r = n/\gcd(p,n)$ and $\mathrm{diag}([\,])$ denotes a diagonal matrix taking the elements of the operand vector as the diagonal entries. From (5.29), (5.30), (5.32) and (5.33), the cyclo spectrum of y is associated with that of x,

$$\begin{aligned}
&\Phi_{yy}(\rho\omega) \\
&= U_{\rho n,1}(\rho\omega)\,\hat{G}\left(e^{j\rho\omega T_s}\right) U^*_{r,p}(p q n\omega) \\
&\quad \cdot \mathrm{diag}\left(\begin{bmatrix} U^*_{p,1}(q n\omega)\,\Phi_{xx}(q n\omega)\,U_{p,1}(q n\omega) \\ U^*_{p,1}(q n\omega + 2\pi/p)\,\Phi_{xx}(q n\omega + 2\pi/p)\,U_{p,1}(q n\omega + 2\pi/p) \\ \vdots \\ U^*_{p,1}(q n\omega + 2\pi(r-1)/p)\,\Phi_{xx}(q n\omega + 2\pi(r-1)/p)\,U_{p,1}(q n\omega + 2\pi(r-1)/p) \end{bmatrix}\right) \\
&\quad U_{r,p}(p q n\omega)\,\hat{G}^*\left(e^{j\rho\omega T_s}\right) U^*_{\rho n,1}(\rho\omega)\,.
\end{aligned} \tag{5.36}$$

Remark: For a fixed frequency ω, either (5.31) or (5.34) can be conceivably computed as the matrix multiplication.

The next example is on the cyclospectrum transformation by LTI systems. The purpose of the example is three-fold: (i) to illustrate what happens beyond the matrix multiplication in (5.31) and (5.34); (ii) to give a concrete example showing a realization of (5.13); (iii) to provide an alternative proof of (5.15).

Example 5.4 Let G be LTI and x be $(\mathrm{CS})_2$, i.e., $m = n = 1$ and $q = p = 2$ in Figure 5.5. Eq. (5.31) becomes

$$\Phi_{yy}(\omega) = U_{21}(\omega)\,\hat{G}\left(e^{j2\omega}\right) U^*_{2,1}(\omega)\,\Phi_{xx}(\omega)\left(U_{21}(\omega)\,\hat{G}\left(e^{j2\omega}\right) U^*_{21}(\omega)\right)^*. \tag{5.35}$$

The unitary matrix $U_{21}(\omega)$ is (see (5.22))

$$U_{21}(\omega) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & e^{-j\omega} \\ 1 & e^{-j(\omega+\pi)} \end{bmatrix}. \tag{5.36}$$

An LTI system is fully characterized by its impulse response $h(\cdot)$, i.e.,

$$y(k) = \sum_{l=-\infty}^{\infty} h(k-l)\,x(l). \tag{5.37}$$

Comparing (5.1) and (5.37) gives the connection between the Green's function and the impulse response

$$g(k,l) = h(k-l). \tag{5.38}$$

The blocked system $\underline{G} = L_2 G L_2^{-1}$ has a transfer matrix (see (5.13))

$$\hat{\underline{G}}(z) = \begin{bmatrix} \hat{G}_{00}(z) & \hat{G}_{01}(z) \\ \hat{G}_{10}(z) & \hat{G}_{11}(z) \end{bmatrix}. \tag{5.39}$$

where (5.20) and (5.38) result

$$\begin{aligned}\hat{G}_{kl}(z) &= \sum_{t=-\infty}^{\infty} g(k+2t,l)\, z^{-t} \\ &= \sum_{t=-\infty}^{\infty} h(2t+k-l)\, z^{-t}.\end{aligned}$$

Thus, (5.39) becomes[5]

$$\underline{\hat{G}}(z) = \begin{bmatrix} \hat{G}_0(z) & z^{-1}\hat{G}_1(z) \\ \hat{G}_1(z) & \hat{G}_0(z) \end{bmatrix}, \tag{5.40}$$

where $\hat{G}_0(z)$ and $\hat{G}_1(z)$ are the well-known type-1 polyphase components of G [147],

$$\begin{aligned}\hat{G}(z) &= h(0) + h(1)\, z^{-1} + h(2)\, z^{-2} + h(3)\, z^{-3} + g(4)\, z^{-4} + g(5)\, z^{-5} + \cdots \\ &= \left(h(0) + h(2)\, z^{-2} + h(4)\, z^{-4} + \cdots\right) + z^{-1}\left(h(1) + h(3)\, z^{-2} + h(5)\, z^{-4} + \cdots\right) \\ &=: \hat{G}_0\left(z^2\right) + z^{-1}\hat{G}_1\left(z^2\right).\end{aligned} \tag{5.41}$$

From (5.36) and (5.40), the product of the first three matrices in (5.35) is

$$\begin{aligned}& U_{2|1}(\omega)\, \underline{\hat{G}}\left(\mathrm{e}^{j2\omega}\right) U_{2|1}^{*}(\omega) \\ &= \frac{1}{2}\begin{bmatrix} 1 & \mathrm{e}^{-j\omega} \\ 1 & \mathrm{e}^{-j(\omega+\pi)} \end{bmatrix}\begin{bmatrix} \hat{G}_0\left(\mathrm{e}^{j2\omega}\right) & \mathrm{e}^{-j2\omega}\hat{G}_1\left(\mathrm{e}^{j2\omega}\right) \\ \hat{G}_1\left(\mathrm{e}^{j2\omega}\right) & \hat{G}_0\left(\mathrm{e}^{j2\omega}\right) \end{bmatrix}\begin{bmatrix} 1 & 1 \\ \mathrm{e}^{j\omega} & \mathrm{e}^{j(\omega+\pi)} \end{bmatrix} \\ &= \begin{bmatrix} \hat{G}_0\left(\mathrm{e}^{j2\omega}\right) + \mathrm{e}^{-j\omega}\hat{G}_1\left(\mathrm{e}^{j2\omega}\right) & 0 \\ 0 & \hat{G}_0\left(\mathrm{e}^{j2\omega}\right) + \mathrm{e}^{-j(\omega+\pi)}\hat{G}_1\left(\mathrm{e}^{j2\omega}\right) \end{bmatrix} \\ &= \begin{bmatrix} \hat{G}\left(\mathrm{e}^{j\omega}\right) & 0 \\ 0 & \hat{G}\left(\mathrm{e}^{j(\omega+\pi)}\right) \end{bmatrix},\end{aligned} \tag{5.42}$$

where the last equality follows from (5.41). Note that the zeros on the off-diagonal entries imply an exact spectrum cancelation. From (5.35), (5.42) and the cyclic spectrum matrix of x (see (5.22))

$$\Phi_{xx}(\omega) = \begin{bmatrix} S_{xx}^{(0)}(\omega) & S_{xx}^{(1)}(\omega) \\ S_{xx}^{(1)}(\omega+\pi) & S_{xx}^{(0)}(\omega+\pi) \end{bmatrix},$$

$\Phi_{yy}(\omega)$ is obtained,

$$\begin{aligned}& \begin{bmatrix} S_{yy}^{(0)}(\omega) & S_{yy}^{(1)}(\omega) \\ S_{yy}^{(1)}(\omega+\pi) & S_{yy}^{(0)}(\omega+\pi) \end{bmatrix} \\ &= \begin{bmatrix} \hat{G}\left(\mathrm{e}^{j\omega}\right) & 0 \\ 0 & \hat{G}\left(\mathrm{e}^{j(\omega+\pi)}\right) \end{bmatrix}\begin{bmatrix} S_{xx}^{(0)}(\omega) & S_{xx}^{(1)}(\omega) \\ S_{xx}^{(1)}(\omega+\pi) & S_{xx}^{(0)}(\omega+\pi) \end{bmatrix}\begin{bmatrix} \hat{G}^{*}\left(\mathrm{e}^{j\omega}\right) & 0 \\ 0 & \hat{G}^{*}\left(\mathrm{e}^{j(\omega+\pi)}\right) \end{bmatrix} \\ &= \begin{bmatrix} \hat{G}\left(\mathrm{e}^{j\omega}\right) S_{xx}^{(0)}(\omega)\, \hat{G}^{*}\left(\mathrm{e}^{j\omega}\right) & \hat{G}\left(\mathrm{e}^{j\omega}\right) S_{xx}^{(1)}(\omega)\, \hat{G}^{*}\left(\mathrm{e}^{j(\omega+\pi)}\right) \\ \hat{G}\left(\mathrm{e}^{j(\omega+\pi)}\right) S_{xx}^{(1)}(\omega+\pi)\, \hat{G}^{*}\left(\mathrm{e}^{j\omega}\right) & \hat{G}\left(\mathrm{e}^{j(\omega+\pi)}\right) S_{xx}^{(0)}(\omega+\pi)\, \hat{G}^{*}\left(\mathrm{e}^{j(\omega+\pi)}\right) \end{bmatrix},\end{aligned}$$

which also proves (5.15). □

[5] Theorem 8.2.1 in [28] gives the generalization of (5.40).

Multirate filter banks are typical examples of LPTV systems and have been of particular interest in signal processing [147]. Unlike the approaches in [113, 99, 112, 79, 2], the statistical properties (in terms of cyclic spectrum) of the output or the reconstructed signal of the filter bank are elegantly found by the blocking technique in the next example.

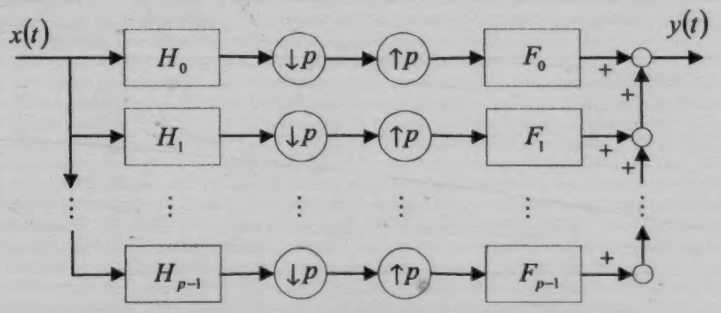

Figure 5.7: A maximally decimated filter bank

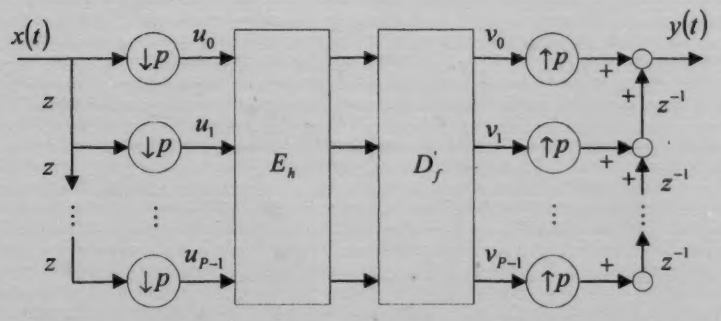

Figure 5.8: Polyphase representation of a filter bank

Example 5.5 Figure 5.7 depicts a maximally decimated filter bank [147], where H_i and F_i $(i = 0, 1, \cdots, p-1)$ are analysis filters and synthesis filters, respectively. Let the input x be $(\mathrm{CS})_p$. With the type-1 polyphase representation of H_k [147], the type-3 polyphase representation of F_k [48], and noble identities [147], Figure 5.7 is equivalent to Figure 5.8 [6]. It is easy to see from (2.8) that $u = [u_0, u_1, \cdots, u_{p-1}]'$ and $v = [v_0, v_1, \cdots, v_{p-1}]'$ in Figure 5.8 are exactly the blocked versions of x and y, respectively. Thus, the original filter bank in Figure 5.7 is represented in terms of the blocked signals u and v with an LTI system $D_f' E_h$. Here both E_h and D_f are $p \times p$ systems; the kl-th elements of their transfer matrices are

$$\left[\hat{E}_h(z)\right]_{kl} = \hat{H}_{kl}^{(p1)}(z), \quad \left[\hat{D}_f(z)\right]_{kl} = \hat{F}_{kl}^{(p3)}(z),$$

[6] A similar observation was also noticed in [30].

where $\hat{H}_{kl}^{(p1)}(z)$ and $\hat{F}_{kl}^{(p3)}(z)$[7] relate to the impulse responses $h_k(\cdot)$ and $f_k(\cdot)$ of H_k and F_k, respectively,

$$\begin{aligned}\hat{H}_{kl}^{(p1)}(z) &= \sum_{t=-\infty}^{\infty} h_k(tp+l)\, z^{-t}, \\ \hat{F}_{kl}^{(p3)}(z) &= \sum_{t=-\infty}^{\infty} f_k(tp-l)\, z^{-t}.\end{aligned}$$

Finally, the cyclic spectrum of the output y can be associated with that of the input x via the following three equations (see (5.31)),

$$\begin{aligned} S_{uu}(p\omega) &= U_{p|1}^{*}(\omega)\, \Phi_{xx}(\omega)\, U_{p|1}(\omega), \\ S_{vv}(\omega) &= \hat{D}_f'\left(e^{j\omega}\right) \hat{E}_h\left(e^{j\omega}\right) S_{uu}(\omega) \left(\hat{D}_f'\left(e^{jp\omega}\right) \hat{E}_h\left(e^{jp\omega}\right)\right)^{*}, \\ \Phi_{yy}(\omega) &= U_{p|1}(\omega)\, S_{vv}(p\omega)\, U_{p|1}^{*}(\omega). \end{aligned}$$

□

5.5 Conclusion

In this paper, we have studied the spectral theory of discrete-time cyclo-stationary signals: the cyclospectrum representation and the cyclospectrum transformation by linear multirate systems. The four types of cyclospectra, namely, the cyclic spectrum, the time frequency representation, the bispectrum and the 2-D spectrum are shown to be closely related and mutually convertible (see Figure 5.3). The cyclospectrum transformation by linear systems are solved in a systematic manner by using multirate systems as the unifying framework and the blocking technique as the main tool. The effects of the blocking operator on cyclo-stationary signals are investigated in Theorem 5.1. The cyclo-stationarity of the output of the multirate system is studied in Theorem 5.2 and the cyclospectrum of the output is associated with that of the input in the form of matrix multiplication in (5.31) and (5.34).

[7] Superscripts $(^{p1})$ and $(^{p3})$ mean the type-1 and type-3 polyphase representations, respectively.

Chapter 6

FIR Modeling for EIV/Closed-Loop Systems

Finite impulse response (FIR) modeling of errors-in-variables (EIV)/closed-loop systems by the traditional correlation analysis usually yields biased estimates due to the additive noises on inputs and outputs. A non-parametric approach, the cyclic correlation analysis (CCRA), provides asymptotically unbiased and consistent estimates. The main feature of the CCRA is to eliminate the adverse effects of stationary noises by exploiting cyclo-stationarity that may exist naturally or be induced artificially. This chapter[1] developes a complete study of the CCRA, including the statistical performance of the estimated FIR model. Frequency-domain expressions of the statistical performance provide guidelines in designing a class of cyclo-stationary signals for modeling. Effectiveness and properties of the CCRA are validated and illustrated by numerical examples.

6.1 Introduction

Consider a setup depicted in Figure 6.1: $G(q)$ is an unknown discrete-time LTI system with impulse response coefficients $g(\cdot)$, i.e.,

$$G(q) := \sum_{l=0}^{\infty} g(l)\, q^{-l};$$

$F_1(q)$ and $F_2(q)$ are some unknown LTI systems in the feedforward and feedback paths, respectively; the input-additive noise $n_u(t)$ perhaps originates from the measurement noise, quantization noise and self-noise [39]; the output-additive noise $n_y(t)$ possibly consists of the measurement noise, disturbance, and model mismatch [85]; $z(t)$ is an external signal to be defined later. The objective[2] is to estimate the first M impulse response coefficients

[1] The chapter has been published in [154, 158].

[2] The objective is very different from the blind identification (e.g., [143]) that only explores output data.

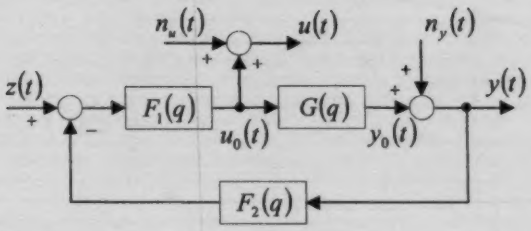

Figure 6.1: A general framework including open-loop, EIV, and closed-loop systems

of $G(q)$, $\{g(0), g(1), \cdots, g(M-1)\}$, from the measured input-output data $\{u(t), y(t)\}_{t=1}^{N}$. We assume:

A6.1 $G(q)$ is asymptotically stable; in addition, if $F_2(q) \neq 0$, the closed-loop system $[1 + F_2(q)F_1(q)G(q)]^{-1}$ is asymptotically stable.

A6.2 Both $n_u(t)$ and $n_y(t)$ are wide-sense stationary, possibly colored and mutually correlated.

A6.3 $n_u(t)$ is uncorrelated with the noise-free input and output, $u_0(t)$ and $y_0(t)$; if $F_2(q) = 0$, so is $n_y(t)$.

The setup in Figure 6.1 is a general framework, and reduces to three common systems after imposing some restrictions: (i) an open-loop system if $F_2(q) = 0$ and $n_u(t)$ is absent, (ii) an (open-loop) EIV system if $F_2(q) = 0$, and (iii) a closed-loop system if $n_u(t)$ is absent.

Estimation of the FIR coefficients of an unknown discrete-time LTI system has found extensive applications in the areas of control and signal processing [22, 77, 122]. First, FIR models are required by some of PID control and model predictive control (MPC) technologies. For example, DMC-plus and RMPCT, the representatives of the fourth generation MPC technologies, estimate the FIR models as the first step upon which low-order parametric models are fixed afterwards [80, 107]. Second, FIR modeling is a simple and effective approach to gain the system knowledge on the variable interaction, dominating time constants, and time delays [19, 49]. Most importantly, the FIR model provides an indispensable comparison with parametric models that have to "guess" the unknown system structure: If the impulse response of a parametric model has good agreement with the FIR model, one would be confident that correct features of the unknown system have been picked up [86]. On the other hand, FIR models are non-parsimonious, i.e., a large number of impulse response coefficients usually have to be estimated; as a result, more data points are needed to reduce the estimation variance [75, 162]. Irrespective of its drawbacks, FIR modeling

has long been advocated as a useful non-parametric analysis accompanying with parametric identification.

The correlation analysis (CRA) is a well-established non-parametric approach to consistently estimate FIR models of open-loop systems without input-additive noises; see details in some books, e.g., [22, 85, 122], and a historical review in [162]. The estimates from the CRA, however, are *biased* for EIV and closed-loop systems.

EIV systems with noise-corrupted inputs and outputs appear in various applications, e.g., system identification, adaptive signal processing, and time series modeling [148]. When only second-order statistics are exploited, a unique solution generally cannot be admitted without imposing additional assumptions [5, 123]. The FIR modeling in [121, 164] requires that $n_u(t)$ and $n_y(t)$ in Figure 6.1 are white noises with known variances or a given ratio between the unknown variances. By contrast, the approaches based on higher-order statistics such as the 3rd-order cross-cumulants [6, 146] can give consistent estimates under the noise assumption A6.2; however, they typically need a large number of data points to achieve estimates with tolerable variances.

Closed-loop systems often arise due to inherent feedback mechanisms, or production, economic and safety reasons [85, 122]. If some external signals outside the feedback loop, like $z(t)$ in Figure 6.1, are available, a so-called joint correlation analysis (JCRA) treats the external signals as instrumental variables, removes the adverse effect of $n_y(t)$ via feedback, and provides asymptotically unbiased FIR models [23, 109]. The main problem of the JCRA is that $z(t)$ may not be strongly correlated with $u(t)$ and $y(t)$, which is mainly determined by closed-loop dynamics; due to the problem, the JCRA may result in unreliable FIR models with large variances (see Example 6.3 later). To alleviate the problem, external signals have to appear at the points as close to system inputs as possible, but the number of such signals would quickly become prohibitive for multivariable systems.

This chapter studies another non-parametric approach, the CCRA, which yields asymptotically unbiased and consistent FIR models for EIV and closed-loop systems. The key characteristic of the CCRA is that the external signal $z(t)$ in Figure 6.1 is assumed to be cyclo-stationary and independent to $n_u(t)$ and $n_y(t)$. Besides being induced by human operations, e.g., amplitude modulation in Section 6.4.1, cyclo-stationarity often exists naturally in industry, e.g., vibrations in rotating machinery [94, 8] and control-loop oscillations [139, 140]. Under these circumstances, exploiting cyclo-stationarity is more reasonable and promising than stalling at stationarity.

Cyclo-stationarity was exploited in estimating the time-difference-of-arrival, namely, an EIV system with only time delay, by a cyclic-correlation-based algorithm [53] that is a special case of the CCRA. Cyclic spectral analysis (CSPA), the frequency-domain counterpart of the CCRA, was proposed in [53, 55] to give asymptotically unbiased frequency-response estimates for EIV systems. The CSPA was generalized for identification of closed-loop systems in [60], where the CCRA was presented without a detailed analysis (Eq. (12) therein and see also [61]). The study of the CSPA was completed in [7] in the sense of developing the statistical performance of the CSPA, including variances of the estimated frequency responses. In [47], frequency responses were estimated from spectral cross-moments and cumulants of high-order cyclo-stationary signals by an algorithm whose computational complexity is comparable to the CSPA.

It is well-known that the CRA and its frequency-domain counterpart, the spectral analysis (SPA), are complementary to each other; so are the CCRA and the CSPA. Hence, this chapter can be regarded as the time-domain counterpart of the work on the CSPA in [53, 55, 7]. In particular, our contribution is three-fold: (i) It provides a relatively complete study of the CCRA in Section 6.3, including its statistical performance, which have not been studied in [53, 60]; (ii) Design a class of cyclo-stationary signals is investigated in Section 6.4 by the aid of analytical results of the statistical performance; (iii) Properties and effectiveness of the CCRA are illustrated and compared with those of the CRA and JCRA via numerical examples in Section 6.5.

6.2 Joint Cyclo-Stationary Signals

This section prepares for the subsequent sections by introducing the concept of joint cyclo-stationary signals and their cyclic correlation and cyclic spectrum. Let us introduce the cyclic correlation/spectrum of jointly $(\mathrm{CS})_p$ signals $x_1(t)$ and $x_2(t)$, whose time-varying cross-correlation $r_{x_1x_2}(t;\tau) := E\{x_1(t)x_2^*(t-\tau)\}$ is periodic in t with period p. Considering the periodicity of $r_{x_1x_2}(t;\tau)$ in t, the cyclic correlation of $x_1(t)$ and $x_2(t)$ is defined as the coefficients of the discrete Fourier series of $r_{x_1x_2}(t;\tau)$ [56, 62],

$$R_{x_1x_2}(k;\tau) = \frac{1}{p}\sum_{t=0}^{p-1} r_{x_1x_2}(t;\tau)\, \mathrm{e}^{-j2\pi kt/p}. \tag{6.1}$$

$R_{x_1x_2}(k;\tau)$ is periodic in k with period p, i.e., $\{R_{x_1x_2}(0;\tau), R_{x_1x_2}(1;\tau), \cdots, R_{x_1x_2}(p-1;\tau)\}$ forms a complete set of cyclic correlations. k is the index of a so-called cycle-frequency domain. Given collected data $\{x_1(t), x_2(t)\}_{t=1}^{N}$, $R_{x_1x_2}(k;\tau)$ can be consistently estimated as

[38, 59, 115],

$$\hat{R}_{x_1x_2}(k;\tau) = \frac{1}{N}\sum_{t=1}^{N-\tau} x_1(t+\tau)\,x_2(t)\,\mathrm{e}^{-j2\pi kt/p},\ \tau \geq 0, \tag{6.2a}$$

$$\hat{R}_{x_1x_2}(k;\tau) = \frac{1}{N}\sum_{t=-\tau}^{N} x_1(t+\tau)\,x_2(t)\,\mathrm{e}^{-j2\pi kt/p},\ \tau < 0. \tag{6.2b}$$

Eq. (6.2) implies that $\hat{R}_{x_1x_2}(0;\tau)$ is the same as the estimate of the stationary cross-correlation $r_{x_1x_2}(\tau) := E\{x_1(t)x_2^*(t-\tau)\}$. The DTFT of $R_{x_1x_2}(k;\tau)$ with respect to τ is named the cyclic spectrum [56, 62],

$$S_{x_1x_2}\left(k;\mathrm{e}^{j\omega}\right) = \sum_{\tau=-\infty}^{\infty} R_{x_1x_2}(k;\tau)\,\mathrm{e}^{-j\omega\tau}. \tag{6.3}$$

$S_{x_1x_2}\left(k;\mathrm{e}^{j\omega}\right)$ inherits the periodicity of $R_{x_1x_2}(k;\tau)$. If $p = 1$, $S_{x_1x_2}\left(0;\mathrm{e}^{j\omega}\right)$ reduces to the stationary (power) spectrum,

$$S_{x_1x_2}\left(\mathrm{e}^{j\omega}\right) := \sum_{\tau=-\infty}^{\infty} r_{x_1x_2}(\tau)\,\mathrm{e}^{-j\omega\tau}.$$

6.3 Cyclic Correlation Analysis

This section presents the definition of the CCRA and the statistical performance of the estimated FIR model. To avoid cumbersome notation, $G(q)$ is assumed to be a SISO system. This is not a restrictive assumption. The FIR model in (6.5) is a linear regression equation without output-relevant terms on the right-hand side of (6.5). Thus, FIR modeling of a multivariable system with m inputs and n outputs is equivalent to modeling of n systems with m inputs and one output, to which the CCRA developed for SISO systems can be extended easily by properly increasing dimensions of matrices and vectors; see e.g., the multivariable CRA in [49, 82]. Note that treating all outputs simultaneously may have *numerical* benefits if impulse responses in some channels are the same [40].

6.3.1 Definition

First of all, we introduce an equality that makes the appearance of the CCRA resemble the CRA: For jointly $(\mathrm{CS})_p$ signals $x_1(t)$ and $x_2(t)$,

$$\bar{E}\left\{x_1(t)\,x_2^*(t-\tau)\,\mathrm{e}^{-j2\pi kt/p}\right\} = R_{x_1x_2}(k;\tau), \tag{6.4}$$

where $k \in [0, p-1]$ (see (4.29)). Based on the equality in (6.4), multiplying both sides of the process-model equation

$$y(t) = \sum_{l=0}^{M-1} g(l)\,[u(t-l) - n_u(t-l)] + n_y(t) \tag{6.5}$$

by $u(t-\tau)e^{-j2\pi kt/p}$ and taking the operation of $\bar{E}\{\cdot\}$ yield

$$R_{yu}(k;\tau) = \sum_{l=0}^{M-1} g(l)\,[R_{uu}(k;\tau-l) - R_{n_u u}(k;\tau-l)] + R_{n_y u}(k;\tau). \tag{6.6}$$

The noise-related terms $R_{n_u u}(k;\tau)$ and $R_{n_y u}(k;\tau)$ in (6.6) are non-zero for $k = 0$, which make the CRA yield biased estimates for closed-loop and EIV systems. On the contrary, $R_{n_u u}(k;\tau)$ and $R_{n_y u}(k;\tau)$ vanish for $k \in [1, p-1]$ and all τ's:

- As the noise-free input $u_0(t)$ is uncorrelated to $n_u(t)$, we have $R_{n_u u}(k;\tau) = R_{n_u u_0}(k;\tau) + R_{n_u n_u}(k;\tau)$, and $R_{n_u u_0}(k;\tau) = 0, \forall\tau$ and $k \in [0, p-1]$. The noise $n_u(t)$ is stationary; thus, (6.1) implies that $R_{n_u n_u}(k;\tau) = 0, \forall\tau$ and $k \in [1, p-1]$. Hence, $R_{n_u u}(k;\tau) = 0, \forall\tau$ and $k \in [1, p-1]$.

- Since $n_y(t)$ and $u_0(t)$ are possibly connected via the LTI system $F_2(q)$ in the feedback path, $R_{n_y u_0}(0;\tau)$ may not vanish, but $R_{n_y u_0}(k;\tau) = 0$ for $k \in [1, p-1]$ and all τ's. Because both $n_u(t)$ and $n_y(t)$ are stationary, $R_{n_y n_u}(k;\tau) = 0$ for $k \in [1, p-1]$ and all τ's. Overall, $R_{n_y u}(k;\tau) = R_{n_y u_0}(k;\tau) + R_{n_y n_u}(k;\tau) = 0$ for $k \in [1, p-1]$ and all τ's.

Thus, (6.6) is simplified to

$$R_{yu}(k;\tau) = \sum_{l=0}^{M-1} g(l)\,R_{uu}(k;\tau-l), \tag{6.7}$$

where $k \in [1, p-1]$, Writing out (6.7) for $\tau = 0, 1, \cdots, M-1$, the impulse response coefficients are estimated as,

$$\begin{bmatrix} \hat{g}(0) & \hat{g}(1) & \cdots & \hat{g}(M-1) \end{bmatrix}' = \begin{bmatrix} \hat{R}_{uu}(k;0) & \hat{R}_{uu}(k;-1) & \cdots & \hat{R}_{uu}(k;1-M) \\ \hat{R}_{uu}(k;1) & \hat{R}_{uu}(k;0) & \cdots & \hat{R}_{uu}(k;2-M) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{uu}(k;M-1) & \hat{R}_{uu}(k;M-2) & \cdots & \hat{R}_{uu}(k;0) \end{bmatrix}^{-1} \begin{bmatrix} \hat{R}_{yu}(k;0) \\ \hat{R}_{yu}(k;1) \\ \vdots \\ \hat{R}_{yu}(k;M-1) \end{bmatrix}. \tag{6.8}$$

Here $\hat{R}_{uu}(k;\tau)$ and $\hat{R}_{yu}(k;\tau)$, obtained from (6.2), are consistent estimators of $R_{uu}(k;\tau)$ and $R_{yu}(k;\tau)$, respectively. Eq. (6.8) can be written in a concise form by introducing some obvious definitions,

$$\hat{\theta} = \left[\Phi_{uu}^{(k)}\right]^{-1} \hat{\eta}_{yu}^{(k)}. \tag{6.9}$$

For a unique solution in (6.9), the square matrix $\Phi_{uu}^{(k)}$ must be nonsingular. This is satisfied by a generalization of the well-known concept of persistent excitation [122, 85]: *A $(CS)_p$ signal $u(t)$ is said to be persistently excited in the cyclo-stationary sense, if its cyclic spectrum $S_{uu}\left(k; e^{j\omega}\right)$ is non-zero for almost all ω and at least one $k \in [1, p-1]$.*

Besides (6.8), an over-determined version of the CCRA can be formed in order to attenuate effects of noises, disturbances and model mismatch,

$$\begin{aligned}
&\left[\begin{array}{cccc} \hat{g}(0) & \hat{g}(1) & \cdots & \hat{g}(M-1) \end{array}\right]' = \\
&\left[\begin{array}{cccc}
\hat{R}_{uu}(k;0) & \hat{R}_{uu}(k;-1) & \cdots & \hat{R}_{uu}(k;1-M) \\
\hat{R}_{uu}(k;1) & \hat{R}_{uu}(k;0) & \cdots & \hat{R}_{uu}(k;2-M) \\
\vdots & \vdots & \ddots & \vdots \\
\hat{R}_{uu}(k;M-1) & \hat{R}_{uu}(k;M-2) & \cdots & \hat{R}_{uu}(k;0) \\
\vdots & \vdots & \ddots & \vdots \\
\hat{R}_{uu}(k;M_0-1) & \hat{R}_{uu}(k;M_0-2) & \cdots & \hat{R}_{uu}(k;M_0-M)
\end{array}\right]^{\dagger}
\left[\begin{array}{c}
\hat{R}_{yu}(k;0) \\
\hat{R}_{yu}(k;1) \\
\vdots \\
\hat{R}_{yu}(k;M-1) \\
\vdots \\
\hat{R}_{yu}(k;M_0-1)
\end{array}\right]
\end{aligned} \tag{6.10}$$

Here $M_0 \geq M$ and the superscript $(^{\dagger})$ denotes the left pseudo-inverse. The matrix inverse in (6.8) or the pseudo-inverse in (6.10) is realized by the QR-method, because a direct computation is sensitive to rounding errors (see e.g., Section 4.5 in [122] and Chapter 5 in [65]).

The CCRA in (6.8) encloses the CRA and JCRA as special cases. Let the setup in Figure 6.1 reduce to a normal open-loop system, i.e., $F_1(q) = 0$, $F_2(q) = 0$, and $n_u(t) = 0$. Eq. (6.8) reduces to the counterpart of the CRA for $k = 0$, because $\hat{R}_{yu}(0;\tau)$ and $\hat{R}_{uu}(0;\tau)$ are the same as the estimators of $r_{yu}(\tau)$ and $r_{uu}(\tau)$, respectively (see Section 6.2). Let us look at the JCRA for closed-loop systems. If $z(t)$ in Figure 6.1 is available, multiplying both sides of (6.5) by $z(t-\tau)e^{-j2\pi kt/p}$ and taking the operation of $\bar{E}\{\cdot\}$ yield

$$R_{yz}(k;\tau) = \sum_{l=0}^{M-1} g(l)\left[R_{uz}(k;\tau-l) - R_{n_u z}(k;\tau-l)\right] + R_{n_y z}(k;\tau). \tag{6.11}$$

Based on (6.11), another version of the CCRA can be obtained analogously to (6.8) by replacing $\hat{R}_{yu}(k;\tau)$ with $\hat{R}_{yz}(k;\tau)$, and $\hat{R}_{uu}(k;\tau)$ with $\hat{R}_{uz}(k;\tau)$. If $k = 0$, this version of the CCRA is the same as the JCRA proposed in [23, 109], as $\hat{R}_{yz}(0;\tau)$ and $\hat{R}_{uz}(0;\tau)$ become the estimators of $r_{yz}(\tau)$ and $r_{uz}(\tau)$, respectively.

6.3.2 Statistical Performance

Substituting the cyclic correlation estimator in (6.2) into (6.8) yields

$$\begin{bmatrix} \hat{g}(0) \\ \hat{g}(1) \\ \vdots \\ \hat{g}(M-1) \end{bmatrix} = \frac{1}{N} \begin{bmatrix} \sum_{t=1}^{N} u^2(t)\, e^{-j2\pi kt/p} \\ \sum_{t=1}^{N-1} u(t+1)\, u(t)\, e^{-j2\pi kt/p} \\ \vdots \\ \sum_{t=1}^{N-M+1} u(t+M-1)\, u(t)\, e^{-j2\pi kt/p} \end{bmatrix}$$

$$\begin{bmatrix} \sum_{t=2}^{N} u(t-1)\, u(t)\, e^{-j2\pi kt/p} & \cdots & \sum_{t=M}^{N} u(t-M+1)\, u(t)\, e^{-j2\pi kt/p} \\ \sum_{t=1}^{N} u^2(t)\, e^{-j2\pi kt/p} & \cdots & \sum_{t=M-1}^{N} u(t-M+2)\, u(t)\, e^{-j2\pi kt/p} \\ \vdots & \ddots & \vdots \\ \sum_{t=1}^{N-M+2} u(t+M-2)\, u(t)\, e^{-j2\pi kt/p} & \cdots & \sum_{t=1}^{N} u^2(t)\, e^{-j2\pi kt/p} \end{bmatrix}^{-1}$$

$$\cdot \frac{1}{N} \begin{bmatrix} \sum_{t=1}^{N} y(t)\, u(t)\, e^{-j2\pi kt/p} \\ \sum_{t=1}^{N-1} y(t+1)\, u(t)\, e^{-j2\pi kt/p} \\ \vdots \\ \sum_{t=1}^{N-M+1} y(t+M-1)\, u(t)\, e^{-j2\pi kt/p} \end{bmatrix}. \tag{6.12}$$

Eq. (6.12) implies that the CCRA is approximately equivalent to the instrumental variable method (IVM) [122, 85],

$$\hat{\theta} = \left[\frac{1}{N} \sum_{t=1}^{N} \zeta(t)\, \varphi'(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^{N} \zeta(t)\, y(t). \tag{6.13}$$

Here $\varphi(t)$ is the regressor of the FIR model in (6.5), i.e.,

$$\varphi'(t) = \begin{bmatrix} u(t) & u(t-1) & \cdots & u(t-M+1) \end{bmatrix}, \tag{6.14}$$

and $\zeta(t)$ is a complex-valued instrumental variable,

$$\zeta'(t) = \begin{bmatrix} u(t)\, e^{-j2\pi kt/p} & u(t-1)\, e^{-j2\pi kt/p} & \cdots & u(t-M+1)\, e^{-j2\pi kt/p} \end{bmatrix}. \tag{6.15}$$

The approximation error arises from the different numbers of data points used in (6.12) and (6.13). For example, the l-th diagonal elements of the square matrices in (6.12) and (6.13), respectively, are $\frac{1}{N}\sum_{t=1}^{N} u^2(t)\, e^{-j2\pi kt/p}$ and $\frac{1}{N}\sum_{t=1}^{N} u^2(t-l)\, e^{-j2\pi kt/p}$. Hence, the approximation error is negligible for $N \gg M$, with the convergence rate of $1/N$.

The approximate equivalence between the CCRA and the IVM implies that the statistical performance of $\hat{\theta}$ in (6.9) could be developed in a manner similar to the asymptotic theory of the IVM; however, special attention has to be paid, since the asymptotic theory of the IVM that holds for stationary signals cannot be in general immediately extended to the CCRA that exploits cyclo-stationarity. In particular, a so-called mixing condition that is exclusive to cyclo-stationary signals has to be satisfied. The mixing condition has several

forms in the literature, e.g., Conditions a)-d) in Proposition 2 in [73], Assumption 1.1 in [38] and Assumptions A-3 and A-4 in [115]; here we adopt the last one as follows:

A6.4 The cyclo-stationary signals $u(t)$ and $y(t)$ satisfy the mixing condition consisting of [115]

$$\sum_{m=0}^{\infty}\sum_{l=m}^{\infty}\left|R_{ab}\left(\alpha;l+\tau\right)R_{cd}\left(\beta;l-\tau\right)\right|<\infty,$$
$$\sum_{l=-\infty}^{\infty}\left|R_{ab^*c^*d}\left(\alpha-\beta;l+\tau_1,l,\tau_2\right)\right|<\infty,$$

for all possible choices of elements $a\left(t\right),b\left(t\right),c\left(t\right)$ and $d\left(t\right)$ from $u\left(t\right),u^*\left(t\right),y\left(t\right)$ and $y^*\left(t\right)$, and all cycle frequencies α and β, and all τ, τ_1, τ_2's. Here the cyclic cumulant is defined as

$$\begin{aligned}R_{abcd}\left(\gamma;\tau_1,\tau_2,\tau_3\right) &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\left[E\left\{a\left(n+\tau_1\right)b\left(n+\tau_2\right)c\left(n+\tau_3\right)d\left(n\right)\right\}\right.\\ &\quad -E\left\{a\left(n+\tau_1\right)b\left(n+\tau_2\right)\right\}E\left\{c\left(n+\tau_3\right)d\left(n\right)\right\}\\ &\quad -E\left\{a\left(n+\tau_1\right)c\left(n+\tau_3\right)\right\}E\left\{b\left(n+\tau_2\right)d\left(n\right)\right\}\\ &\quad \left.-E\left\{a\left(n+\tau_1\right)d\left(n\right)\right\}E\left\{b\left(n+\tau_2\right)c\left(n+\tau_3\right)\right\}\right]\mathrm{e}^{-j2\pi\gamma n}.\end{aligned}$$

Loosely speaking, the mixing condition in A6.4 says that the statistical dependence between any two samples of $u(t)$ and $y(t)$ decays rapidly enough as the temporal separation between them increases, and $u(t)$ and $y(t)$ have finite fourth-order cumulants.

Proposition 6.1 *Under assumptions A6.1-A6.4 and $u(t)$ being persistently excited in the cyclo-stationary sense, $\hat{\theta}$ in (6.9) is the asymptotically unbiased and consistent estimate of the true impulse response coefficients $\theta := \left[\begin{array}{cccc} g\left(0\right) & g\left(1\right) & \cdots & g\left(M-1\right)\end{array}\right]'$; in particular, $\hat{\theta}$ is asymptotically normal-distributed (AsN) with zero mean and variance $P_{\hat{\theta}}/N$,*

$$\sqrt{N}\left(\hat{\theta}-\theta\right)\to AsN\left(0,P_{\hat{\theta}}\right),\ \text{as } N\to\infty, \tag{6.16a}$$

where

$$P_{\hat{\theta}}=\lambda\left[\bar{E}\left\{\zeta\left(t\right)\varphi'\left(t\right)\right\}\right]^{-1}\left[\bar{E}\left\{\zeta_F\left(t\right)\zeta_F^*\left(t\right)\right\}\right]\left\{\left[\bar{E}\left\{\zeta\left(t\right)\varphi'\left(t\right)\right\}\right]^{-1}\right\}'. \tag{6.16b}$$

Here $\varphi\left(t\right)$ and $\zeta\left(t\right)$ are given in (6.14) and (6.15), respectively. The vector $\zeta_F\left(t\right)$ and the noise variance λ are defined in the following proof.

Proof of Proposition 6.1: As $N \to \infty$, the CCRA estimate in (6.9) is equivalent to $\hat{\theta}$ in (6.13). From (6.5) and (6.13),

$$\sqrt{N}\left(\hat{\theta}-\theta\right) \to \left[\frac{1}{N}\sum_{t=1}^{N}\zeta(t)\varphi'(t)\right]^{-1}\frac{1}{\sqrt{N}}\sum_{t=1}^{N}\zeta(t)\varepsilon(t), \tag{6.17}$$

where

$$\varepsilon(t) := n_y(t) - G(q)\, n_u(t).$$

Assumptions A6.1, A6.2 and A6.3 imply that the residual $\varepsilon(t)$ is stationary and uncorrelated with $\zeta(t)$ in (6.15) for $k \in [1, p-1]$ (see the paragraph between (6.6) and (6.7)). From the Wold representation, the stationary signal $\varepsilon(t)$ can be modeled as the output of an LTI system $H(q)$[3] driven by white noise with variance λ. Under assumption A6.4, $\zeta(t)$ and $\varepsilon(t)$ are stochastic signals with decaying dependence and finite fourth-order cumulants. As a result, we have [84]

$$\frac{1}{\sqrt{N}}\sum_{t=1}^{N}\zeta(t)\varepsilon(t) \to AsN(0,P), \tag{6.18}$$

where

$$\begin{aligned} P &= \lim_{N\to\infty} E\left\{\zeta(t)\varepsilon(t)\varepsilon^*(t)\zeta^*(t)\right\} \\ &= \lambda \bar{E}\left\{\zeta_F(t)\zeta_F^*(t)\right\}. \end{aligned}$$

Here $\zeta_F(t)$ is obtained by filtering $\zeta(t)$ in (6.15) through $H\left(q^{-1}\right)$, i.e., $\zeta_F(t) = H\left(q^{-1}\right)\zeta(t)$. The convergence in distribution (6.16) follows directly from (6.17) and (6.18). □

Propositions 6.2 and 6.3 constitute the frequency-domain counterpart of Proposition 6.1. However, (6.19) and (6.22) cannot be obtained directly from the results of the IVM, e.g., (8.102) and (9.84) in [85], in spite of the resembling forms, because their proofs require several results involving cyclic correlation/spectra instead of (stationary) correlation/spectra.

Proposition 6.2 *The convergence of the estimate from the CCRA is asymptotically characterized in the frequency domain as*

$$\hat{\theta} \to sol_{\bar{\theta}}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[G\left(e^{j\omega},\theta\right) - G\left(e^{j\omega},\bar{\theta}\right)\right] W_1\left(e^{-j\omega}\right) S_{uu}\left(k;e^{j\omega}\right) d\omega = 0\right\}, \tag{6.19}$$

with $W_1\left(e^{j\omega}\right) = \left[\begin{array}{cccc} 1 & e^{j\omega} & \cdots & e^{j(M-1)\omega} \end{array}\right]'$.

Proof of Proposition 6.2: The CCRA is approximately equivalent to the IVM; thus, $\hat{\theta}$ convergences to the solution to the function $\bar{E}\left\{\zeta(t)\left[y(t) - G\left(q,\bar{\theta}\right)u(t)\right]\right\} = 0$ as $N \to \infty$,

[3] To avoid nonlinear optimization, $H(q)$ is usually confined to a high-order AR model.

i.e.,

$$\theta \to \arg\min_{\theta} \left[E\left\{ C(\theta) \left[y(t) - G(q,\theta) u(t) \right] \right\} = 0 \right]$$

With $C(t)$ as (6.15), the k-th element of the vector $E\left\{ C(\theta) \left[y(t) - G(q,\theta) u(t) \right] \right\}$ is

$$\begin{aligned} & \left[E\left\{ C(\theta) \left[y(t) - G(q,\theta) u(t) \right] \right\} \right]_k \\ &= E\left\{ u(t-k) e^{-j\omega_k t} \cdot \left[G(q,\theta)(u)(t) - u_0(t) + u_0(t) - G(q,\theta) u(t) \right] \right\} \\ &= E\left\{ u(t-k) e^{-j\omega_k t} \cdot \left[G(q,\theta) - G(q,\theta) \right] u(t) \right\} \\ &= R_{ue}(k,t), \end{aligned}$$

where the last equality comes from (6.4) and a definition, $e(t) = \left[G(q,\theta) - G(q,\theta) \right] u(t)$. The cyclic spectrum between $e(t)$ and $u(t)$ is [10, 68],

$$S_{ue}(k,e^{j\omega}) = \left[G(e^{j\omega},\theta) - G(e^{j\omega},\theta) \right] S_{uu}(k,e^{j\omega}). \tag{6.30}$$

Eq. (6.13) is obtained from (6.28) and the inverse version of (6.8),

$$R_{ue}(k,\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{ue}(k,e^{j\omega}) e^{j\omega\tau} d\omega. \tag{6.31}$$

□

Proposition 6.5 *The asymptotic covariance of the estimate from the CVEM has a frequency domain expression:*

$$Cov\left\{ \theta \right\} \to \frac{1}{N} P^{-1} Q \left(P^{-1} \right)^T, \tag{6.32}$$

where

$$\begin{aligned} P &= \frac{1}{2\pi} \int_{-\pi}^{\pi} W_k\left(e^{j(\omega_k - \omega)t} \right) S_{uu}(k,e^{j\omega}) d\omega, \\ Q &= \frac{1}{2\pi} \int_{-\pi}^{\pi} W_k\left(e^{j\omega t} \right) H\left(e^{-j\omega} \right) S_{uu}(k,e^{j\omega}) H^H\left(e^{j\omega} \right) d\omega, \end{aligned}$$

with

$$W_k\left(e^{j\omega t} \right) = \begin{bmatrix} 1 & e^{-j\omega} & \cdots & e^{-j(M-1)\omega} \\ e^{j\omega} & 1 & \cdots & e^{-j(M-2)\omega} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(M-1)\omega} & e^{j(M-2)\omega} & \cdots & 1 \end{bmatrix}$$

Proof of Proposition 6.5 The objective is to find the frequency-domain expression of (6.16b). The m-th row and n-th column element of the matrix $E\left\{ C(t) u^T(t) \right\}$ is,

$$\begin{aligned} & \left[E\left\{ C(t) u^T(t) \right\} \right]_{m,n} \\ &= E\left\{ u(t) \cdot m(t) e^{-j(m-n)\omega_k t} u_0(t-n) \right\} \\ &= R_{uu}(k, m-n) e^{j(m-n)\omega_k (t-n)} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{uu}(k,e^{j\omega}) e^{j(m-n)(\omega - \omega_k)(t-n)} d\omega. \end{aligned} \tag{6.33}$$

where the second and last equalities are from (6.4) and (6.21), respectively. The m-th row and n-th column element of the matrix $\bar{E}\left\{\zeta_F(t)\,\zeta_F^*(t)\right\}$ is,

$$\begin{aligned}
&\left[\bar{E}\left\{\zeta_F(t)\,\zeta_F^*(t)\right\}\right]_{m,n} \\
&= \bar{E}\left\{H\left(q^{-1}\right)u(t-m)\,\mathrm{e}^{-j2\pi kt/p}H^*\left(q^{-1}\right)u^*(t-n)\,\mathrm{e}^{j2\pi kt/p}\right\} \\
&= \bar{E}\left\{v(l)\,v^*(l-m+n)\right\} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} S_{vv}\left(0;\mathrm{e}^{j\omega}\right)\mathrm{e}^{j\omega(m-n)}\mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} H\left(\mathrm{e}^{-j\omega}\right)S_{uu}\left(0;\mathrm{e}^{j\omega}\right)H^*\left(\mathrm{e}^{-j\omega}\right)\mathrm{e}^{j(m-n)\omega}\mathrm{d}\omega. \qquad (6.24)
\end{aligned}$$

Here $v(t) := H\left(q^{-1}\right)u(t)$, and the last equality is obtained from the cyclic-spectrum transformation relationship [56],

$$S_{vv}\left(k;\mathrm{e}^{j\omega}\right) = H\left(\mathrm{e}^{-j\omega}\right)S_{uu}\left(k;\mathrm{e}^{j\omega}\right)H^*\left(\mathrm{e}^{-j(\omega+2\pi k/p)}\right).$$

Substituting (6.23) and (6.24) into (6.16) gives (6.22). □

6.4 Discussion

Two important elements of the CCRA need to be determined, namely, the cyclo-period p and the cycle frequency k. The discussion is classified into two situations according to whether cyclo-stationarity is induced artificially or exists naturally.

6.4.1 Induced Cyclo-Stationarity

There are several artificial operations to generate cyclo-stationary signals, such as amplitude modulation, time-index modulation, multirate sampling, and multirate filtering [61]. The massive possibilities in these operations imply that it is difficult to have a general solution of designing cyclo-stationary signals for modeling. Here, we focus on a class of amplitude modulation signals and select its configurations based on the statistical performance of the CCRA developed in Section 6.3.2.

The class of amplitude modulation signals is

$$z(t) = \cos\left(\frac{\pi t}{p}\right)s(t), \qquad (6.25)$$

where $s(t)$ is a zero-mean stationary signal to be defined. $z(t)$ is $(\mathrm{CS})_p$, for its time-varying correlation is

$$r_{zz}(t;\tau) = \frac{1}{2}\left[\cos\left(\frac{\pi\tau}{p}\right) + \cos\left(\frac{2\pi t}{p} - \frac{\pi\tau}{p}\right)\right]r_{ss}(\tau).$$

$z(t)$ and its relatives, e.g., $z(t) = \cos(\omega_0 t)s(t)$, have been studied in [55, 94, 61, 145]; however, these studies did not discuss how to choose p or ω_0.

A good index in determining p and k is the degree of cyclo-stationarity (DCS) that measures the distance between stationarity and cyclo-stationarity. Several definitions of the DCS were proposed in [166]; one of them is

$$\mathrm{DCS}_{uu}^{k} = \frac{\int_{-\pi}^{\pi} \left| S_{uu}\left(k; \mathrm{e}^{j\omega}\right)\right| \mathrm{d}\omega}{\int_{-\pi}^{\pi} \left| S_{uu}\left(0; \mathrm{e}^{j\omega}\right)\right| \mathrm{d}\omega},$$

which is a ratio of the energy in the k-th cycle-frequency domain to the energy of stationary components. DCS_{uu}^{k} matches the signal-to-noise ratio in Proposition 6.3, and is suitable in this context. A property of the DCS is [166],

$$\mathrm{DCS}_{uu}^{k} \leq 1. \tag{6.26}$$

It is desired to select p and k ($k \in [1, p-1]$) so that DCS_{uu}^{k} or DCS_{zz}^{k} is as close to 1 as possible. The cyclic spectrum of $z(t)$ in (6.25) is

$$\begin{aligned} S_{zz}\left(k; \mathrm{e}^{j\omega}\right) &= \frac{1}{4} S_{ss}\left(\mathrm{e}^{j\left(\omega+\frac{\pi}{p}\right)}\right) \delta\left(k-1\right) + \frac{1}{4} S_{ss}\left(\mathrm{e}^{j\left(\omega-\frac{\pi}{p}\right)}\right) \delta\left(k-p+1\right) \\ &\quad + \frac{1}{4}\left[S_{ss}\left(\mathrm{e}^{j\left(\omega+\frac{\pi}{p}\right)}\right) + S_{ss}\left(\mathrm{e}^{j\left(\omega-\frac{\pi}{p}\right)}\right)\right] \delta\left(k\right). \end{aligned} \tag{6.27}$$

Here $\delta(\cdot)$ is Kronecker's delta function,

$$\delta(l) = \begin{cases} 1, & l = 0, \\ 0, & \text{else.} \end{cases}$$

If $p = 2$, $S_{zz}\left(0; \mathrm{e}^{j\omega}\right)$ and $S_{zz}\left(1; \mathrm{e}^{j\omega}\right)$ are the same, and the maximum of the DCS_{zz}^{k} is achieved, i.e., $\mathrm{DCS}_{zz}^{1} = 1$; if $p \geq 3$, only $S_{zz}\left(1; \mathrm{e}^{j\omega}\right)$ and $S_{zz}\left(p-1; \mathrm{e}^{j\omega}\right)$ are non-zero, and contain at most the half energy of $S_{zz}\left(0; \mathrm{e}^{j\omega}\right)$, i.e., $0 \leq \mathrm{DCS}_{zz}^{k} \leq 0.5$.

If $p \geq 3$, the multiple estimates obtained from the CCRA for different k's, denoted as $\hat{\theta}_k = [\hat{g}_k(0), \hat{g}_k(1), \cdots, \hat{g}_k(M-1)]'$, could be combined to reach an estimate with smaller variance. Since $\hat{\theta}_k$ is normal-distributed (see Proposition 6.1), the weighted linear combination

$$\hat{g}(l)_{opt} = \sum_{k=1}^{p-1} w_k(l) \hat{g}_k(l), \quad l = 0, 1, \cdots, M-1.$$

achieves the minimum variance [51]. It is known from *Lemma 2* in [115] that $\hat{R}_{uu}(k; \tau)$'s for different k and τ are correlated to each other; as a result, $\hat{g}_k(l)$'s are correlated, and computation of the optimal weight $w_k(l)$ requires the unknown correlation coefficients among

$\hat{g}_k(l)$'s. Alternatively, we propose an aggregated CCRA,

$$\hat{\theta} = \begin{bmatrix} \Phi_{uu}^{(1)} \\ \Phi_{uu}^{(2)} \\ \vdots \\ \Phi_{uu}^{(p-1)} \end{bmatrix}^{\dagger} \begin{bmatrix} \eta_{yu}^{(1)} \\ \eta_{yu}^{(2)} \\ \vdots \\ \eta_{yu}^{(p-1)} \end{bmatrix}. \tag{6.28}$$

Proposition 6.1 is applicable to $\hat{\theta}$ in (6.28) by using a different instrument variable,

$$\tilde{\zeta}(t) = \left[\begin{array}{cccc} \zeta_1'(t) & \zeta_2'(t) & \cdots & \zeta_{p-1}'(t) \end{array} \right],$$

where $\zeta_k'(t)$ is given in (6.15). Eq. (6.28) is solved by the QR-method, and encloses the CCRA in (6.9) as a special case. The estimated impulse response coefficients from the aggregated CCRA have the variances no larger than the estimates from the CCRA exploiting one single k. In addition, Propositions 6.2 and 6.3 imply that the improvement achieved by the aggregated CCRA depends on the cyclic spectra of $u(t)$; see Example 6.2 in Section 6.5. For $z(t)$ in (6.25), only two estimates of the CCRA for $k = 1$ and $k = p - 1$ are to be aggregated. Due to the possible correlation between $R_{uu}(1;\tau)$ and $R_{uu}(p-1;\tau)$, or $S_{uu}(1;e^{j\omega})$ and $S_{uu}(p-1;e^{j\omega})$, the aggregation at best has the effect of decreasing the variances by half. Therefore, we select $p = 2$ and $k = 1$ as final choices, since $\mathrm{DCS}_{zz}^{k} \leq 0.5$ for $p \geq 3$ and $\mathrm{DCS}_{zz}^{1} = 1$ for $p = 2$.

6.4.2 Existing Cyclo-Stationarity

When cyclo-stationarity exists naturally in some instances, the CCRA becomes a natural choice. Vibration signals in rotating machinery can be modeled in the form of (see e.g., Eq. (9) in [7])

$$x(t) = \sum_{i} A_i(t)\, e^{j\omega_i t}.$$

Under certain conditions, $A_i(t)$ is stationary and ω_i's are in harmonic, so that the vibration signal $x(t)$ is cyclo-stationary. The auto-covariance function of an oscillating signal is oscillatory with the same period as the oscillation in the time trend [139, 140]; thus, the control-loop oscillations are conjectured to be cyclo-stationary. If the cyclo-period p is not known *a priori*, it can be estimated by one of cyclo-period estimation methods in Chapter 3. If $p \geq 3$, the cycle frequency k should be chosen after estimating $S_{uu}(k;e^{j\omega})$ by smoothed cyclic periodograms or correlograms (see e.g., Eq. (17.22) in [61]), and the aggregated CCRA in (6.28) can be used for those k's with non-zero DCS_{uu}^{k}. Applying the CCRA in these applied situations is one of our ongoing studies.

6.5 Numerical Illustration

This section presents three numerical examples to illustrate the properties and effectiveness of the proposed CCRA. First, Example 6.1 shows that the CCRA yields asymptotically *unbiased* estimates for EIV systems, while estimates from the CRA are biased due to the input-additive noise. The asymptotic variance in (6.16) is consistent with the result obtained from the multiple Monte Carlo simulations in Example 6.1.

Example 6.1 An EIV system depicted in Figure 6.1 has the configuration:

$$G(q) = \frac{0.09516q^{-3}}{1 - 0.9048q^{-1}}, \quad F_1(q) = 1, \quad F_2(q) = 0.$$

$n_u(t)$ is white noise with zero mean and variance 0.2, abbreviated as WN(0, 0.2); $n_y(t)$ is the output of an LTI filter

$$H_y(q) = \frac{0.125 + 0.033q^{-1}}{1 - 0.3679q^{-1}}$$

driven by $n_u(t)$; thus, $n_y(t)$ and $n_u(t)$ are mutually correlated. $z(t)$ is the same as that in (6.25) with $p = 2$, i.e., $z(t) = \cos\left(\frac{\pi t}{2}\right) s(t)$. Here $s(t)$ is a stationary random-binary sequence (RBS) with normalized frequency band $[0, 1]$ and values $\{1, -1\}$. The input and output signal-to-noise ratios (SNR) are defined as,

$$\text{SNR}_i = \sqrt{\frac{\frac{1}{N}\sum_{t=0}^{N-1} E\left\{u_0^2(t)\right\}}{E\left\{n_u^2(t)\right\}}}, \quad \text{SNR}_0 = \sqrt{\frac{\frac{1}{N}\sum_{t=0}^{N-1} E\left\{y_0^2(t)\right\}}{E\left\{n_y^2(t)\right\}}}.$$

In this example, $\text{SNR}_i \approx 1.58$ and $\text{SNR}_o \approx 2.32$.

100 Monte Carlo simulations are performed with different realizations of $n_u(t)$ and $z(t)$. FIR models are estimated from the CCRA ($k = 1$ in (6.8)) and CRA with the measured input-output data $\{u(t), y(t)\}_{t=1}^{2000}$. The sample mean of FIR models from the CCRA and that from the CRA are shown in Figures 6.2 and 6.3, respectively. Clearly, the CCRA yields unbiased estimates of impulse response coefficients, while the estimates from the CRA are biased due to $n_u(t)$. The sample mean of the asymptotic variances calculated from (6.16) is presented in the form of 3-standard deviation (dash lines in Figure 6.2), and is consistent with the sample variances of the FIR models obtained from the 100 simulations, also shown in the form of 3-standard deviation (plus lines in Figure 6.2) — see Appendix B.9 in [122] for the accuracy of Monte Carlo analysis. The 3-sample time delay of $G(q)$ is correctly estimated by looking at the first three impulse response coefficients and their asymptotic variances in Figure 6.2

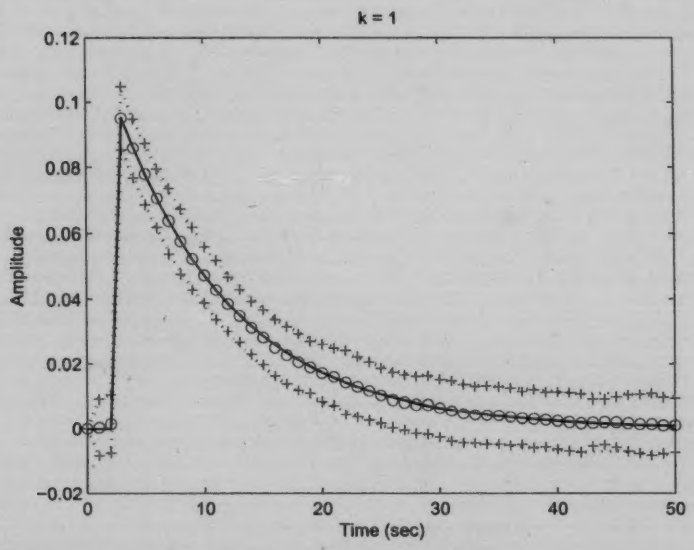

Figure 6.2: FIR modeling of an EIV system from the CCRA ($N = 2000$): the true impulse response coefficients (solid) and the estimates from the CCRA (circle) with the 3-standard deviation band from (6.16) (dash) and that from 100 Monte Carlo simulations (plus).

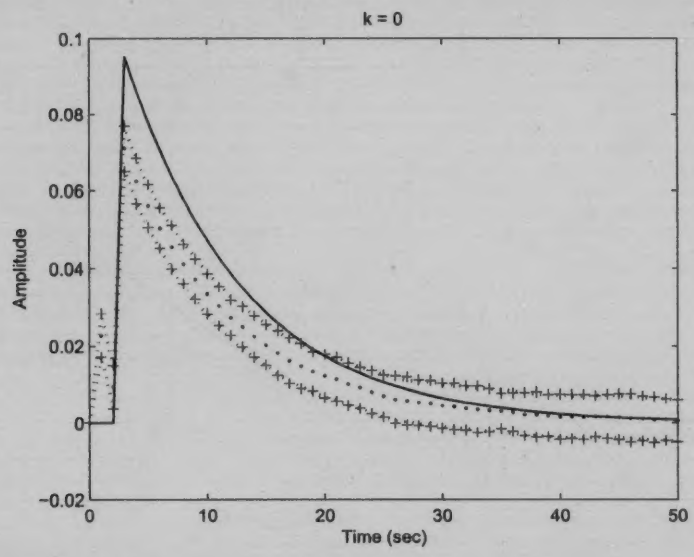

Figure 6.3: FIR modeling of an EIV system from the CRA ($N = 2000$): the true impulse response coefficients (solid) and the estimates from the CRA (dot) with two 3-standard deviation band from (6.16) (dash) and that from 100 Monte Carlo simulations (plus).

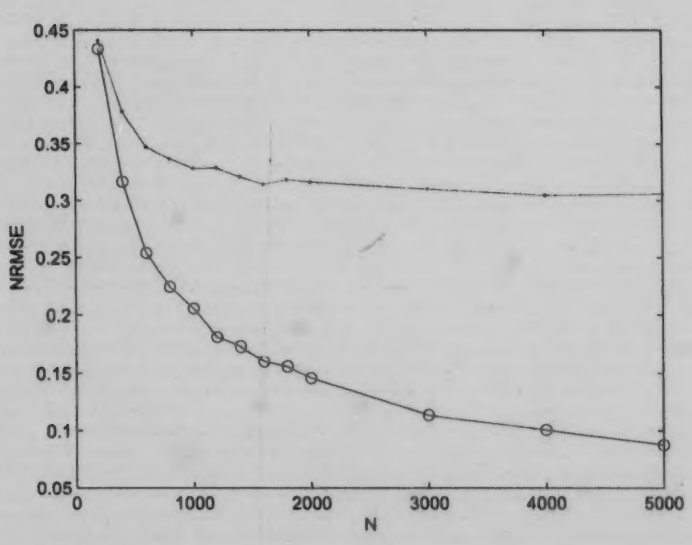

Figure 6.4: NRMSEs of the CCRA (circle) and CRA (dot) for different N

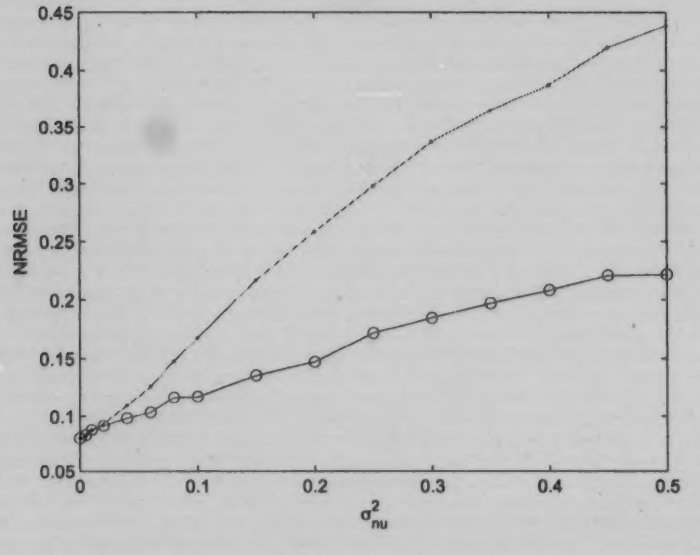

Figure 6.5: NRMSEs of the CCRA (circle) and CRA (dot) for different noise levels

Both the CCRA and CRA require sufficient data points to exhibit their asymptotic properties. A normalized root-mean-square-error (NRMSE) is adopted as a measurement of the identification performance,

$$\text{NRMSE} = \frac{1}{\|g(l)\|_2}\sqrt{\frac{1}{M}\sum_{l=0}^{M-1}\|\hat{g}(l) - g(l)\|_2^2}$$

Figure 6.4 presents the sample means of the NRMSEs from the CCRA and CRA as a function of the data length N. 100 Monte Carlo simulations are performed for each N. A moderate value of N, e.g., 1000 or 2000, is sufficient to eliminate the adverse effects of $n_u(t)$ and $n_y(t)$.

Proposition 6.3 and (6.26) imply that the estimate from the CCRA, denoted as $\hat{\theta}_k$, generally has the variance greater or equal to that from the CRA, denoted as $\hat{\theta}_0$. This is due to a fact that the noise component of $u(t)$ contributes in reducing the variance of $\hat{\theta}_0$, and has no such a good effect on $\hat{\theta}_k$:

$$\begin{aligned}\text{Cov}\left\{\hat{\theta}_0\right\} &\propto \int_{-\pi}^{\pi}\frac{1}{|S_{u_0u_0}(0;e^{j\omega}) + S_{n_un_u}(e^{j\omega})|}\mathrm{d}\omega,\\ \text{Cov}\left\{\hat{\theta}_k\right\} &\propto \int_{-\pi}^{\pi}\frac{1}{|S_{u_0u_0}(k;e^{j\omega})|}\mathrm{d}\omega\end{aligned}$$

Note that the noise component of $u(t)$ in a closed-loop system may come from $n_y(t)$ via feedback. Figure 6.5 assesses the CRA and CCRA in terms of the NRMSE for different noise levels, i.e., the variance of $n_u(t)$, denoted as $\sigma_{n_u}^2$, varies from 0 to 0.5. For each noise level, 100 Monte Carlo simulations are performed with fixed $N = 2000$; to see the effect of $n_u(t)$ only, the noise source of $n_y(t)$ is fixed to be WN(0, 0.2) and is independent of $n_u(t)$. The CRA yields the estimate with smaller NRMSE than the CCRA only at a very low noise level, e.g., $\text{NRMSE}\left\{\hat{\theta}_0\right\} = 0.0831$ and $\text{NRMSE}\left\{\hat{\theta}_1\right\} = 0.0837$ at $\sigma_{n_u}^2 = 0.005$. However, the CCRA quickly outperforms the CRA in terms of smaller NRMSE as the noise level increases. □

In Example 6.1, the cyclo-period $p = 2$ and the CCRA in (6.8) is exploited with $k = 1$. The selection of p is based on the discussion in Section 6.4.1; Example 6.2 is to confirm the discussion by comparing the aggregated CCRA in (6.28) and the CCRA with one single k.

Example 6.2 Let us look at the EIV system in Example 1 under the same configuration generating Figures 6.2 and 6.3 with one modification: the cyclo-period p of $z(t)$ in (6.25) varies from 2 to 7. For each cyclo-period, 100 Monte Carlo simulations are performed. Tables 6.1 lists the sample mean of the NRMSEs of $\hat{\theta}_1$ from the CCRA with $k = 1$, and of

$[0,1]$	NRMSE$\left\{\hat{\theta}_1\right\}$	NRMSE$\left\{\hat{\theta}_{1,p-1}\right\}$
$p=2$	0.1398	0.1398
$p=3$	0.2137	0.2111
$p=4$	0.2006	0.1988
$p=5$	0.2031	0.2010
$p=6$	0.2060	0.2024
$p=7$	0.2073	0.2039

Table 6.1: NRMSEs of $\hat{\theta}_1$ (CCRA) and $\hat{\theta}_{1,p-1}$ (aggregated CCRA) for $s(t)$ with frequency-band $[0,1]$

$[0.2,0.8]$	NRMSE$\left\{\hat{\theta}_1\right\}$	NRMSE$\left\{\hat{\theta}_{1,p-1}\right\}$
$p=2$	0.1543	0.1543
$p=3$	0.2108	0.1882
$p=4$	0.2283	0.1808
$p=5$	0.2482	0.2028
$p=6$	0.3087	0.2723
$p=7$	0.3574	0.3333

Table 6.2: NRMSEs of $\hat{\theta}_1$ (CCRA) and $\hat{\theta}_{1,p-1}$ (aggregated CCRA) for $s(t)$ with frequency-band $[0.2,0.8]$

$\hat{\theta}_{1,p-1}$ from the aggregated CCRA with $k=1$ and $k=p-1$. The aggregated CCRA has little improvement over the CCRA, because $S_{uu}\left(1;e^{j\omega}\right)$ and $S_{uu}\left(p-1;e^{j\omega}\right)$ are the same for white noise $s(t)$ in (6.25) (see (6.27)). Table 6.2 presents another group of results with a different $s(t)$ in (6.25): the frequency-band of $s(t)$ is $[0.2,0.8]$. In this case, $S_{uu}\left(1;e^{j\omega}\right)$ and $S_{uu}\left(p-1;e^{j\omega}\right)$ have different magnitude distributions. The aggregated CCRA always performs better than the CCRA; the improvement is more significant than that in Table 6.1. In Table 6.1 or 6.2, the case of $p=2$ achieves the smallest NRMSE, which is consistent with the discussion in Section 6.4.1. □

Finally, Example 6.3 illustrates that both CCRA and JCRA give unbiased estimates for closed-loop systems, whereas the estimates from the CRA are biased due to the effect of output-additive noise via feedback. Moreover, the CCRA does not suffer from the main problem of the JCRA (see Section 6.1) and provides reliable estimates.

Example 6.3 A testing closed-loop system in [161] is adopted here (see Figure 6.1):

$$G(q)=\frac{q^{-1}+0.7q^{-2}}{1+1.4q^{-1}+0.45q^{-2}},\quad F_1(q)=0.33+0.033q^{-1}-0.4q^{-2},\quad F_2(q)=1.$$

Here $n_u(t)$ is absent; $n_y(t)$ is WN$(0,0.05)$; $z(t)$ is the same as that in Example 6.1; $\mathrm{SNR}_o\approx 2.40$.

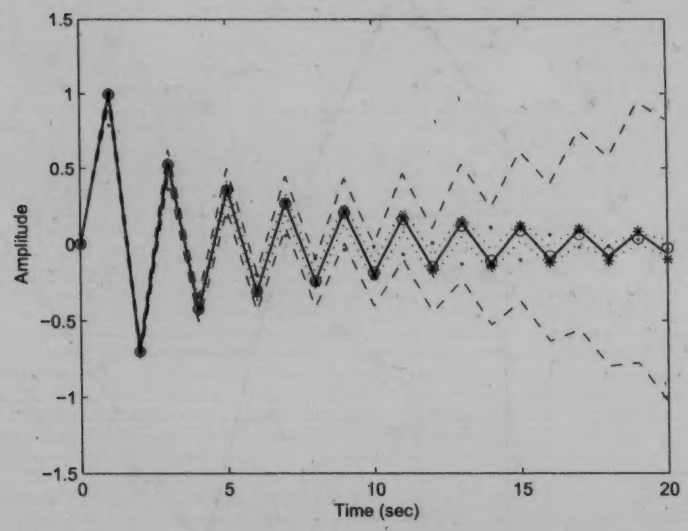

Figure 6.6: FIR modeling of a closed-loop system ($N = 2000$): the true impulse response coefficients (solid), the estimates from the CRA (dot), from the CCRA (circle) with 3-standard deviation band (short-dash), from the JCRA (star) with 3-standard deviation band (long-dash).

We perform 100 Monte Carlo simulations with different realizations of $n_y(t)$ and $z(t)$, and estimate FIR models from the CCRA ($k = 1$ in (6.8)), CRA and JCRA. Besides $\{u(t), y(t)\}_{t=1}^{2000}$, $\{z(t)\}_{t=1}^{2000}$ has to be assumed available for the JCRA. As expected, the CRA results in a biased FIR model (dots in Figure 6.6) due to the correlation between $n_y(t)$ and $u(t)$ via feedback; on the contrary, the CCRA and JCRA both give asymptotically unbiased estimates (circles and stars in Figure 6.6). However, the FIR model from the JCRA is not reliable due to the large variances (long-dash lines in Figure 6.6) obtained from 100 Monte Carlo simulations, in contrast to the CCRA (short-dash lines in Figure 6.6). The condition number of the corresponding matrix in the JCRA is at the mercy of the closed-loop dynamics, and has a possibility of being ill-conditioned. In this example, the averaged condition number in the CCRA is 23.65, while that in the JCRA is 139.57! □

6.6 Conclusion

This chapter studies a non-parametric identification approach, the CCRA, to estimate asymptotically consistent and unbiased FIR models for EIV systems and closed-loop systems by exploiting cyclo-stationarity. The CCRA is studied in a complete manner, including its

statistical performance in Propositions 6.1-6.3. We believe that the proposed CCRA would be a useful non-parametric technique for its effectiveness and properties illustrated by the numerical examples in Section 6.5.

One of the important open problems is to design cyclo-stationary input signals. In this paper, a class of cyclo-stationarity signals, $z(t)$ in (6.25), is investigated. In [7], another class of amplitude modulation signals, namely, a periodic rectangular sequence $a(t) = a(t+T)$ modulating white noise, was studied with a partial solution: the ratio of the burst-time to the period T was optimized to minimize the variances of estimated frequency responses, whereas the important parameter T was left to user's decision. There may be other cyclo-stationary signals that lead to better properties or new features in certain sense for system modeling. This certainly deserves attention in the future study.

Chapter 7

Blind Identification of Hammerstein Systems

This chapter[1] proposes a new blind approach to identification of Hammerstein systems, where a static nonlinearity precedes a linear dynamic system. By exploiting input's piecewise constant property, the parameters of the linear dynamics are consistently estimated from the information of the output only, after which the unmeasurable inner signal is uniquely reconstructed. The noise effect is explicitly considered in both the parameter and inner signal estimation. The estimation of the system orders and time delay are studied on the basis of two groups of basic equations obtained by polyphase decomposition. Magneto-rheological (MR) dampers are semi-active control devices to reduce vibrations of various dynamic structures. By designing a real-time laboratory experiment, we apply the proposed blind approach and build a Hammerstein model for MR dampers.

7.1 Introduction

Hammerstein systems form a class of block-oriented nonlinear models, where a static nonlinearity precedes a linear dynamic subsystem. Many real-time processes can be well represented by Hammerstein models, such as heat exchangers [46], electrical drives [16], thermal microsystems [134], physiological systems [43] and sticky valves [127].

Identification of Hammerstein systems can be classified according to whether the nonlinearity and linear system are identified together or separately. In the former class, there are many existing methods, e.g., iterative methods [98, 129, 165, 43] and non-iterative two-stage methods [27, 130, 150, 9, 63, 64] — see e.g., Chapter 5 in [74] for a recent overview of these methods. The later class basically has three groups of approaches, namely, the

[1]The chapter has been published in [159, 157, 160].

correlation stochastic approaches, the relay feedback approaches and the blind approaches. A major distinction among them is the assumption on input properties. The correlation stochastic approaches [18, 66, 102] require white Gaussian inputs to isolate the nonlinearity and estimate impulse response coefficients of the linear system first. The relay feedback approaches [4, 88, 16, 134, 101, 11] introduce binary-valued inputs by user's design in open-loop systems or by relay feedback-controllers in closed-loop systems; under binary-valued inputs, identification of the linear system is decoupled from that of the nonlinearity. The blind approaches [131, 14] aim at the main difficulty in identification of Hammerstein systems: the inner signal between the nonlinearity and linear system is unmeasurable. Inputs are assumed to be piece-wise constant for certain consecutive samples, based on which the linear system is estimated separately from identification of the nonlinearity.

The blind approaches are very useful for the case where the structure of the nonlinearity is unknown, because they visualize the shape of the nonlinearity by estimating the unmeasurable inner signal beforehand and avoid a wild guess on the structure. Such a case arises when the nonlinearity has many possible structures or is hard to be represented by parametric models. In particular, one real-time application is to capture the nonlinearities of actuators in feedback control systems. It has been found that control valves account for about one third of control-loop oscillations [17, 44]. The nonlinearity of an actuator has a variety of possible structures, e.g., deadband, saturation, backlash and hysteresis [33]. Srinivasan *et al.* [127] demonstrated the potentiality of exploiting the Hammerstein identification in diagnosing valve stiction; however, their approach was based on a separable least-squares identification algorithm proposed in [10] and applicable to only the nonlinearity with known structure and one single unknown parameter.

The first contribution of this chapter is to propose a new blind approach to identification of Hammerstein systems. The new approach has two main differences with the existing blind approaches [131, 14]. (i) The noise-corrupted cases are considered instead of the noise-free ones in [14]. The battle against noises leads to a new series of static errors-in-variables (EIV) systems. By contrast to [131], the realization of static EIV systems significantly reduces the complexity of estimating the numerator parameters (Section 7.3). (ii) The intermediate signal was estimated in [14] by passing output measurements through an inverse[2]of the identified linear system; its drawback is the propagation of the output noise into the estimates. The counterpart in [131] is unnecessary complicated due to the way in estimating the numerator parameters. We estimate the inner signal differently by a

[2]A special treatment is necessary if the linear system is non-minimum phase [11].

least-squares method borrowed from the blind equalization; by doing so, the noise effect is reduced. On the other hand, the proposed blind approach inherits some technical features from those in [131, 14], e.g., Assumptions A7.1 and A7.5 in Section 7.2 and the estimation of the denominator parameters in Section 7.3.2.

The second contribution lies at modeling of magneto-rheological (MR) dampers in a real-time experiment. MR dampers are semi-active control devices to reduce vibrations of various dynamic structures. MR fluids, whose viscosities vary with input voltages/currents, are exploited in providing controllable damping forces. MR dampers were first introduced by B.F. Spencer to civil applications in mid 1990s. In 2001, MR dampers were applied to the cable-stayed Dongting Lake Bridge in China and the National Museum of Emerging Science and Innovation Building in Japan, which are the world's first full-scale implementations in civil structures [31]. Modeling of MR dampers has received considerable attention [126, 163, 32]. Recently, Song, Ahmadian & Southward [124] proposed a nonparametric model that becomes a Hammerstein system if the input current/voltage is constant. We will design a real-time identification experiment for MR dampers and build a Hammerstein model by the proposed blind approach.

The rest of the paper is organized as follows. Section 7.2 describes the problem and gives some necessary assumptions. Section 7.3 estimates the parameters of the linear system as well as the system orders and time delay. With the identified linear dynamics, the inner signal is estimated in Section 7.4. The consistency of parameter estimation and the uniqueness of inner signal estimation are proved in Section 7.5. Section 7.6 illustrates the proposed blind approach by a simulated numerical example. Modeling of MR dampers by the proposed blind approach is presented in Section 7.7, followed by concluding remarks in Section 7.8.

7.2 Problem Description

Consider a discrete-time Hammerstein system with sampling period h depicted in Figure 7.1. Our objective is to identify an LTI causal dynamic system $G(q)$ and a static nonlinearity $f(\cdot)$ from the measured input $u(t)$ and measured output $y(t)$ that is contaminated by colored noise $v(t)$. The inner signal $x(t)$ is unmeasurable — the main difficulty in identification of Hammerstein systems.

We make the following assumptions throughout the paper:

A7.1 The input $u(t)$ is piece-wise constant for p consecutive samples. Because the nonlin-

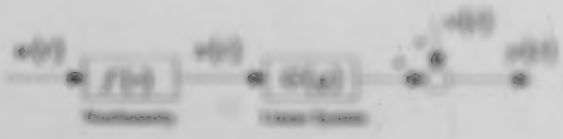

Figure 7.1: A discrete-time Hammerstein system with sampling period k.

verify f is static[3], the inner signal $v(t)$ inherits the same property (i.e.,

$$v(t) = v(t-1) = 0, \text{ for } (kp+1) \le t \le (kp+p-1), \forall k \in \mathbb{Z}_+. \tag{7.1}$$

A7.2 The input $u(t)$ is persistently excited and has more than two different values. In addition, $u(t)$ may be correlated to the noise $v(t-d)$ for $d \ge 1$, for instance, if a feedback loop exists between $y(t)$ and $u(t)$.

A7.3 The linear system and noise dynamics can be described by an autoregressive with exogenous variables (ARX) model,

$$y(t) = \frac{B(q)}{A(q)} u(t-\tau) + \frac{1}{A(q)} v(t), \tag{7.2}$$

where

$$\begin{aligned} A(q) &= 1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_{n_a} q^{-n_a}, \\ B(q) &= b_1 q^{-1} + b_2 q^{-2} + \cdots + b_{n_b} q^{-n_b}. \end{aligned}$$

Here the noise source $v(t)$ is white with zero mean and variance σ^2.

A7.4 If the time delay is decomposed as $\tau = kp + \tau_0$ for $k \in \mathbb{Z}_+$ and $\tau_0 \in \{0, p\}$, then k is known a priori.

A7.5 The upper bound n_a^0 of the autoregressive order n_a is known a priori and p is no less than $(n_a^0 + 1)$, i.e., $p \ge n_a^0 + 1$.

A7.6 $B(q)$ does not have a zero at 1, i.e., $\sum_{i=1}^{n_b} b_i \neq 0$.

Assumption A7.1 is satisfied in several scenarios. It combines two prices from user's design, e.g., the modeling experiments for MR dampers in Section 7.7. Another scenario occurs in sampled-data systems depicted in Figure 7.3. The output is sampled with period h, p times faster than the input updating period $T = ph$. Because of the ZOH, a fast-rate

[3] In fact, f does not have to be static in order to pass the piece-wise constant property of $u(t)$ to $v(t)$, e.g., the backlash nonlinearity in Example 7.1.

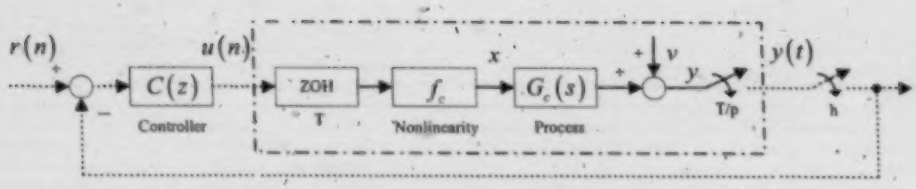

Figure 7.2: A sampled-data Hammerstein system

input with sampling period h is available by interpolation and has the piece-wise constant property in (7.1).

Assumption A7.2 is a standard identifiability condition. For Hammerstein systems, "identifiability" is understood with a gain ambiguity between $f(\cdot)$ and $G(q)$; the ambiguity can be removed by letting $b_1 = 1$. Assumption A7.3 is based on a well-known fact that a high-order ARX model is capable of approximating any linear system arbitrarily well (Page 336 in [85]). Assumption A7.4 essentially arises from a character of the blind identification that the information of output only cannot distinguish time delays $\tau_1 = k_1 p$ and $\tau_2 = k_2 p$ for $k_1 \neq k_2$ (to be clarified in Section 7.3.4). We assume $\tau \in [0, p)$ in the sequel without loss of generality, because the known portion of τ can be removed by shifting output data. Assumption A7.5 is inherent in the existing blind approaches as well; in fact, Theorem 2.1 in [15] says that Assumption A7.5 is a sufficient and necessary condition for $G(q)$ to be blindly identifiable. Assumption A7.6 is a mild assumption satisfied by many systems.

7.3 Identification of Linear Dynamics

Two groups of equations are obtained by the polyphase decomposition of involved signals. Based on them, the parameters in the linear system, the orders n_a, n_b and the time delay τ are estimated.

7.3.1 Two Groups of Basic Equations

For the time being, the orders n_a and n_b and the time delay τ are assumed to be known; thus, τ becomes zero after shifting data properly. We will return to the estimation of n_a, n_b and τ later in Section 7.3.4. By denoting $w(t) := A(q)\, y(t)$, (7.2) becomes

$$\begin{aligned} w(t) &= B(q)x(t) + e(t) \\ &= \sum_{j=1}^{n_b} b_j x(t-j) + e(t). \end{aligned} \tag{7.3}$$

Subtracting two consecutive samples $w(t)$ and $w(t-1)$ yields

$$w(t) - w(t-1) = \sum_{j=1}^{n_b} b_j \left(x(t-j) - x(t-j-1) \right) + e(t) - e(t-1). \tag{7.4}$$

Define the difference signals $\Delta_w(t) := w(t) - w(t-1)$, and $\Delta_x(t)$ and $\Delta_e(t)$ likewise. Eq. (7.4) can be rewritten as

$$\Delta_w(t) = \sum_{j=1}^{n_b} b_j \Delta_x(t-j) + \Delta_e(t),$$

whose z-transformation is

$$\Delta_w(z) = \sum_{j=1}^{n_b} b_j z^{-j} \Delta_x(z) + \Delta_e(z). \tag{7.5}$$

The polyphase decomposition of $\Delta_w(z)$ for the factor of p is [147, 48]

$$\begin{aligned}
\Delta_w(z) &= \sum_t \Delta_w(t) z^{-t} \\
&=: \sum_k \sum_{l=1}^{p} \Delta_w(kp+l) z^{-(kp+l)} \\
&= \sum_{l=1}^{p} z^{-l} \sum_k \Delta_w(kp+l) (z^p)^{-k} \\
&=: \sum_{l=1}^{p} z^{-l} \Delta_w^{(l)}(z^p).
\end{aligned} \tag{7.6}$$

Similarly, the polyphase decomposition of $\Delta_x(z)$ is

$$\Delta_x(z) = \sum_{l=1}^{p} z^{-l} \Delta_x^{(l)}(z^p), \tag{7.7}$$

where

$$\Delta_x^{(l)}(z^p) := \sum_k \Delta_x(kp+l) (z^p)^{-l}.$$

Thanks to the property of $x(t)$ in (7.1), $\Delta_x(t)$ is nonzero only at $t = kp$, $\forall k \in \mathbb{Z}_+$. Thus, we have

$$\begin{aligned}
\Delta_x^{(l)}(z^p) &= 0, \quad l = 1, 2, \cdots, p-1, \\
\Delta_x^{(p)}(z^p) &\neq 0.
\end{aligned}$$

Eq. (7.7) reduces to

$$\Delta_x(z) = z^{-p} \Delta_x^{(p)}(z^p). \tag{7.8}$$

Substituting (7.6) and (7.8) into (7.5) yields

$$\sum_{l=1}^{p} z^{-l}\Delta_w^{(l)}(z^p) = \sum_{l=1}^{n_b} b_l z^{-l} z^{-p}\Delta_x^{(p)}(z^p) + \sum_{l=1}^{p} z^{-l}\Delta_e^{(l)}(z^p). \tag{7.9}$$

With Assumption A7.5, i.e., $p \geq (n_b^0 + 1) > n_b$, (7.9) implies

$$\begin{aligned}
\Delta_w^{(n_b+1)}(z^p) &= \Delta_e^{(n_b+1)}(z^p), && (7.10a)\\
\Delta_w^{(n_b+2)}(z^p) &= \Delta_e^{(n_b+2)}(z^p), && (7.10b)\\
&\vdots \\
\Delta_w^{(p)}(z^p) &= \Delta_e^{(p)}(z^p), && (7.10c)
\end{aligned}$$

and

$$\begin{aligned}
\Delta_w^{(1)}(z^p) &= b_1 z^{-p}\Delta_x^{(p)}(z^p) + \Delta_e^{(1)}(z^p), && (7.11a)\\
\Delta_w^{(2)}(z^p) &= b_2 z^{-p}\Delta_x^{(p)}(z^p) + \Delta_e^{(2)}(z^p), && (7.11b)\\
&\vdots \\
\Delta_w^{(n_b)}(z^p) &= b_{n_b} z^{-p}\Delta_x^{(p)}(z^p) + \Delta_e^{(n_b)}(z^p). && (7.11c)
\end{aligned}$$

The two groups of equations (7.10) and (7.11) are the bases to estimate the parameters in $A(q)$ and $B(q)$. Moreover, they make the estimation of the numerator order n_b and the time delay τ possible. The idea of exploiting the polyphase decomposition is inspired by [13] where only noise-free cases were considered. A timing diagram of polyphase decomposition of signals is available in many textbooks, e.g., Figure 1.4 in [48]. The two groups of equations (7.10) and (7.11) can also be seen from Figure 7.3 appeared later.

7.3.2 Estimation of $A(q)$

In the time domain, (7.10) implies

$$\Delta_w(kp+l) = \Delta_e(kp+l), \quad l = n_b+1, n_b+2, \cdots, p, \ \forall k \in \mathbb{Z}_+. \tag{7.12}$$

Define Δ_y analogously to Δ_w and Δ_e, i.e., $\Delta_y(kp+l) := y(kp+l) - y(kp+l-1)$. Eq. (7.12) is written in terms of Δ_y as

$$\begin{aligned}
\Delta_y(kp+l) &= \sum_{i=1}^{n_a} -a_i\Delta_y(kp+l-i) + \Delta_e(kp+l) \\
&= \phi_y'(k)\,\theta_a + \Delta_e(kp+l), \quad l = n_b+1, n_b+2, \cdots, p, \ \forall k \in \mathbb{Z}_+, \quad (7.13)
\end{aligned}$$

where

$$\theta_a = [\, a_1 \;\; a_2 \;\; \cdots \;\; a_{n_a} \,]',$$
$$\phi_y(k) = [\, -\Delta_y(kp+l-1) \;\; -\Delta_y(kp+l-2) \;\; \cdots \;\; -\Delta_y(kp+l-n_a) \,]'.$$

Eq. (7.13) is linear in the parameter a_i. However, if the ordinary least-squares method (LSM) is applied to (7.13) with the collected data $\{y(t)\}_{t=1}^{N}$, i.e.,

$$\bar{\theta}_a = \left[\frac{1}{K}\sum_{k=0}^{K-1}\phi_y(k)\,\phi_y'(k)\right]^{-1}\frac{1}{K}\sum_{k=0}^{K-1}\phi_y(k)\,\Delta_y(kp+l),$$

the resulted estimate $\bar{\theta}_a$ is biased. Here K is the largest integer less than or equal to N/p. The bias arise from the correlation between the noise term $\Delta_e(kp+l)$ with the first regressor $\Delta_y(kp+l-1)$, which can be resolved by a bias-compensated LSM. The difference between $\bar{\theta}_a$ and θ_a is

$$\bar{\theta}_a - \theta_a = \left[\frac{1}{K}\sum_{k=0}^{K-1}\phi_y(k)\,\phi_y'(k)\right]^{-1}\frac{1}{K}\sum_{k=0}^{K-1}\phi_y(k)\,\Delta_e(kp+l).$$

As $K \to \infty$,

$$\lim_{K\to\infty}\frac{1}{K}\sum_{k=0}^{K-1}\phi_y(k)\,\Delta_e(kp+l) = E\{\phi_y(k)\,\Delta_e(kp+l)\} = [\, \sigma^2 \;\; 0 \;\; \cdots \;\; 0 \,]'_{1\times n_a}.$$

Recall that σ^2 is the variance of the noise source $e(t)$. Thus, θ_a can be estimated without bias by explicitly compensating the noise effect, i.e.,

$$\hat{\theta}_a^{(l)} = \left[\frac{1}{K}\sum_{k=0}^{K-1}\phi_y(k)\,\phi_y'(k)\right]^{-1}\left\{\left[\frac{1}{K}\sum_{k=0}^{K-1}\phi_y(k)\,\Delta_y(kp+l)\right] - \begin{bmatrix}\hat{\sigma}^2\\0\\\vdots\\0\end{bmatrix}\right\}. \tag{7.14}$$

A consistent estimate of σ^2 was developed in Theorem 1 of [76]: $\hat{\sigma}^2$ is the smaller one of the roots (x_1, x_2) of a quadratic equation in x, i.e.,

$$\hat{\sigma}^2 = \min(x_1, x_2). \tag{7.15}$$

The quadratic equation in x is

$$0.5 g_{11} x^2 - x + \bar{\sigma}^2 = 0,$$

where

$$g_{11} = \left\{\left[\frac{1}{K}\sum_{k=0}^{K-1}\phi_y(k)\,\phi_y'(k)\right]^{-1}\right\}_{11},$$
$$\bar{\sigma}^2 = \frac{1}{K}\sum_{k=0}^{K-1}\left(\Delta_y(kp+l) - \phi_y'(k)\,\bar{\theta}_a\right)^2.$$

Here $\{\cdot\}_{11}$ stands for the first-column and first-row element of the operand matrix.

7.3.3 Estimation of $B(q)$

Let $b_1 = 1$ to remove a scalar ambiguity between the process $G(q)$ and the nonlinearity f; thus, (7.11a) gives

$$z^{-p}\Delta_w^{(p)}(z^p) = \Delta_w^{(1)}(z^p) - \Delta_e^{(1)}(z^p). \tag{7.16}$$

Substituting (7.16) into the other equations in (7.11) yields

$$\Delta_w^{(l)}(z^p) = b_l\left(\Delta_w^{(1)}(z^p) - \Delta_e^{(1)}(z^p)\right) + \Delta_e^{(l)}(z^p), \quad l = 2, 3, \cdots, n_b,$$

which implies that in the time domain,

$$\Delta_w(kp+l) = b_l\left(\Delta_w(kp+1) - \Delta_e(kp+1)\right) + \Delta_e(kp+l), \quad l = 2, 3, \cdots, n_b, \ \forall k \in \mathbb{Z}_+. \tag{7.17}$$

Taking $\Delta_w(kp+1)$ and $\Delta_w(kp+l)$ as the input and output, respectively, (7.17) is a static EIV system with the input and output noises, $\Delta_e(kp+1)$ and $\Delta_e(kp+l)$, respectively. It is straightforward to derive the following properties of $\Delta_e(kp+1)$ and $\Delta_e(kp+l)$ by considering the facts that $p \geq 2$ and $e(t)$ is white noise with variance σ^2:

1. Both $\Delta_e(kp+1)$ and $\Delta_e(kp+l)$ are white noises, having the same variance $2\sigma^2$.
2. If $l = 2$, $\Delta_e(kp+1)$ and $\Delta_e(kp+l)$ are correlated; their correlation is equal to $\sigma^2\delta(k)$, where $\delta(\cdot)$ denotes the Dirac delta function.
3. If $l = 3, 4, \cdots, n_b$, $\Delta_e(kp+1)$ and $\Delta_e(kp+l)$ are mutually independent.

Due to the second property, some of the existing identification methods for EIV systems, e.g., the total least-squares method, cannot be applied directly to estimate b_2. In parallel to Section 7.3.2, we propose a new bias-compensated LSM to estimate b_l,

$$\hat{b}_l = \left[\frac{1}{K}\sum_{k=0}^{K-1}\Delta_w^2(kp+1) - 2\hat{\sigma}^2\right]^{-1}\left[\frac{1}{K}\sum_{k=0}^{K-1}\Delta_w(kp+1)\Delta_w(kp+l) + \hat{\sigma}^2\delta(l-2)\right]. \tag{7.18}$$

Here $\hat{\sigma}^2$ has been obtained in (7.15). Under Assumptions A7.1-A7.5, Theorem 7.1 in Section 7.5 proves that $\hat{b}_l$ in (7.18) is a consistent estimate.

7.3.4 Determination of n_a, n_b and τ

This subsection briefly describes the principles to determine the orders n_a, n_b and the time delay τ; these principles are implemented in a slightly complicated way in Section 7.6 and are also illustrated by a simulation example therein and another real-time example in Section 7.7.3.

As $\hat{\theta}_a^{(l)}$ in (7.14) is based on the linear regression in (7.13), the determination of n_a is rather standard by the model structure determination methods in Section 11.5 in [122] and Section 16.4 in [85]. Here the so-called Akaike information criterion (AIC) is adopted,

$$V^{(l)}(n_a) = \left(1 + 2\frac{n_a}{K}\right)\frac{1}{K}\sum_{k=0}^{K-1}\left(\Delta_y^{(l)}(k) - \left[\phi_y^{(l)}(k)\right]'\hat{\theta}_a^{(l)}\right)^2 \qquad (7.19)$$

for $l \in [n_b + 1,\ p]$. Thus, n_a would be the integer associated with the minimum value of $V^{(l)}(\hat{n}_a)$.

The time delay τ is determined by a careful observation of the two groups of equations in (7.10) and (7.11). If τ is nonzero, to let the counterpart of (7.4) reach one of the equations in (7.10) needs two inequalities: $kp \le t - \tau - n_b - 1$ and $t - \tau - 1 \le kp + p - 1$ for some integer $k \in \mathbb{Z}_+$. The inequalities say that two time delays $\tau_1 = k_1 p + \tau_0$ and $\tau_2 = k_2 p + \tau_0$ for $k_1 \ne k_2$ and $\tau_0 \in [0, p)$ cannot be distinguished from the information of $y(t)$ only. Hence, the time delay τ in (7.2) is assumed in the range of $[0, p)$ without loss of generality under Assumption A7.4. In this case, (7.10) and (7.11) respectively hold for

$$n_b + 1 + \tau \le l \le p + \tau \text{ and } 1 + \tau \le l \le n_b + \tau.$$

They imply that $V^{(p)}(n_a)$ defined in (7.19) has a non-zero contribution from Δu extra to that from Δe for n_b times, and a sole contribution from Δe for $(p - n_b)$ times, when the output is consecutively shifted forward by $l = 0, 1, \cdots, p-1$ samples, i.e., $y(t) = y(t + l)$. Therefore, $\hat{\tau}$ is the largest number of the consecutive shifts resulting $(p - n_b)$ equivalent smallest numbers among all the p values of $V^{(p)}(n_a)$. As n_b is unknown in practice, we may choose $l = p$. Then, n_a and τ would be the integers associated with the minimum value of $V^{(p)}(\hat{n}_a)$'s among different combinations of $\hat{n}_a$ and $\hat{\tau}$ — see the examples in Sections 7.6 and 7.7.3.

The two groups of equations in (7.10) and (7.11) also tell the order n_b once n_a, τ and $A(q)$ have been estimated. Due to the contribution from Δu, the first n_b $\Delta_w^{(l)}$'s have the larger variances than the rest, $\Delta_w^{(n_b+1)}$, $\Delta_w^{(n_b+2)}$, $\cdots$, $\Delta_w^{(p)}$, which have the same variance $2\sigma^2$; thus, n_b is determined as the difference between p and the number of equivalent smallest $\Delta_w^{(l)}$'s.

In fact, n_b can also be determined from the number of equivalent $V^{(p)}(n_a)$'s; however, we would prefer to decouple the estimation of n_a and τ from that of n_b, which is found in simulations to be easier and more robust.

7.4 Inner Signal Estimation

With $\hat{A}(q)$ and $\hat{B}(q)$ in hand, the second step of the blind Hammerstein identification is to estimate the unmeasurable inner signal $x(t)$, or equivalently its slow-rate version $X(n) := x(pn)$. We first connects the special FIR system in (7.3) with its equivalent single-input and multiple-output (SIMO) counterpart, and then estimate $X(n)$ by a method borrowed from the blind equalization (see e.g., [1]). The method has no differentiation on minimum or non-minimum phase systems.

In general, a fast-rate FIR model like (7.3) with the sampling period h can be described as

$$w(t) = \sum_{j=1}^{\infty} h(j)\, x(t-j) + e(t). \tag{7.20}$$

Owing to the property in (7.1), (7.20) is equivalent to a slow-rate SIMO FIR model with the sampling period T,

$$W(n) = \sum_{k=0}^{\infty} H(k)\, X(n-k) + E(n), \tag{7.21}$$

where

$$\begin{aligned} W(n) &= \begin{bmatrix} w_1(n) \\ w_2(n) \\ \vdots \\ w_p(n) \end{bmatrix} := \begin{bmatrix} w(pn+1) \\ w(pn+2) \\ \vdots \\ w(pn+p) \end{bmatrix}, \\ E(n) &= \begin{bmatrix} e_1(n) \\ e_2(n) \\ \vdots \\ e_p(n) \end{bmatrix} := \begin{bmatrix} e(pn+1) \\ e(pn+2) \\ \vdots \\ e(pn+p) \end{bmatrix}, \\ X(n) &:= x(pn). \end{aligned}$$

By substituting (7.20) into (7.21) and exploiting the property in (7.1), the impulse response of the SIMO model is connected with that of the fast-rate model in (7.20) as

$$H(k) = \begin{bmatrix} h_1(k) \\ h_2(k) \\ \vdots \\ h_p(k) \end{bmatrix} = \begin{bmatrix} \sum_{l=0}^{p-1} h(kp+1-l) \\ \sum_{l=0}^{p-1} h(kp+2-l) \\ \vdots \\ \sum_{l=0}^{p-1} h(kp+p-l) \end{bmatrix}. \tag{7.22}$$

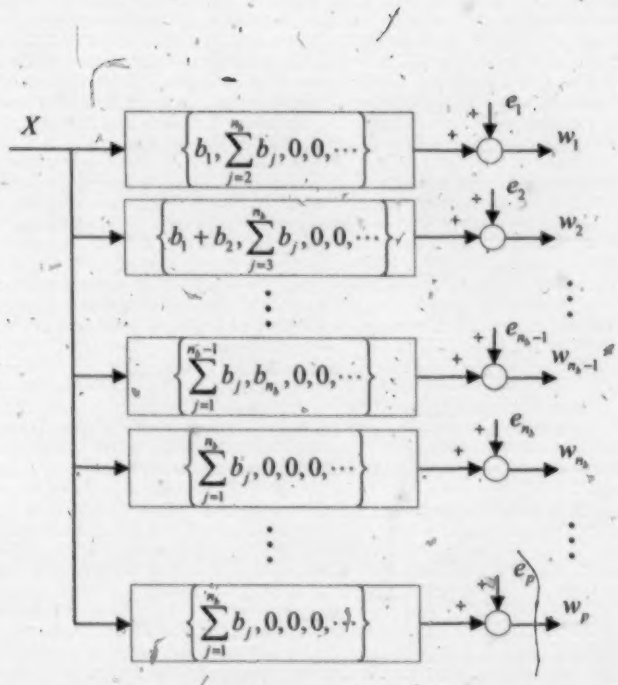

Figure 7.3: An equivalent slow-rate SIMO FIR model

We now return to the special fast-rate FIR model in (7.3) that has the impulse responses,

$$h(0) = 0, h(1) \triangleq b_1, h(2) = b_2, \cdots, h(n_b) = b_{n_b}, h(n_{b+1}) = 0, \cdots, h(p) = 0, \cdots$$

Its equivalent slow-rate SIMO model according to (7.22) has only two non-zero impulse responses, i.e.,

$$H(0) = \begin{bmatrix} h_1(0) \\ h_2(0) \\ \vdots \\ h_{n_b}(0) \\ h_{n_b+1}(0) \\ \vdots \\ h_p(0) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_1 + b_2 \\ \vdots \\ \sum_{j=1}^{n_b} b_j \\ \sum_{j=1}^{n_b} b_j \\ \vdots \\ \sum_{j=1}^{n_b} b_j \end{bmatrix}, \; H(1) = \begin{bmatrix} h_1(1) \\ h_2(1) \\ \vdots \\ h_{n_b}(1) \\ h_{n_b+1}(1) \\ \vdots \\ h_p(1) \end{bmatrix} = \begin{bmatrix} \sum_{j=2}^{n_b} b_j \\ \sum_{j=3}^{n_b} b_j \\ \vdots \\ b_{n_b} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (7.23)$$

and $H(k) = 0_{p\times 1}, \forall k \geq 2$. For clarity, the SIMO model is depicted in Figure 7.3. The data of the i-th output $w_i(n)$ are associated with those of the unknown input $X(n)$ as

$$\begin{aligned} \mathbf{w}_i &= [\, w_i(1) \;\; w_i(2) \;\; \cdots \;\; w_i(K-1) \,]' \\ &= [\, w(p+i) \;\; w(2p+i) \;\; \cdots \;\; w((K-1)p+i) \,]' \end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} h_i(1) & h_i(0) & 0 & \cdots & 0 \\ 0 & h_i(1) & h_i(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_i(1) & h_i(0) \end{bmatrix} \begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(K-1) \end{bmatrix} + \begin{bmatrix} e_i(0) \\ e_i(1) \\ \vdots \\ e_i(K-1) \end{bmatrix} \\
&=: \mathbf{H}_i \mathbf{X} + \mathbf{E}_i,
\end{aligned}
$$

where $h_i(0)$ and $h_i(1)$ are given in (7.23). Since $e_1, e_2, \cdots, e_p$ are mutually independent and have the same variance, it is reasonable to stack all $\mathbf{y}_i$'s together, i.e.,

$$
\mathbf{W} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_p \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_p \end{bmatrix} \mathbf{X} + \begin{bmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \\ \vdots \\ \mathbf{E}_p \end{bmatrix} =: \mathbf{H}\mathbf{X} + \mathbf{E}. \tag{7.24}
$$

Based on (7.24) with $w(t) = \hat{A}(q)y(t)$ and $\hat{b}_l$, a least-squares estimate is obtained,

$$
\hat{\mathbf{X}} = \left(\hat{\mathbf{H}}'\hat{\mathbf{H}}\right)^{-1} \hat{\mathbf{H}}'\mathbf{W}, \tag{7.25}
$$

which is also a maximum-likelihood estimate if e_l is white and Gaussian noise.

7.5 Theoretical Analysis

This section analyzes the consistency of the estimated parameters $\hat{a}_i, \hat{b}_j$ and the uniqueness of the inner signal estimation in Sections 7.3 and 7.4, respectively.

Lemma 7.1 *Under Assumptions A7.1-A7.5, the matrix*

$$
\lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \phi_y(k)\, \phi_y'(k)
$$

is positive definite.

Proof of Lemma 7.1: It follows with some modifications from Lemma 1 in [131] □

Theorem 7.1 *Under Assumptions A7.1-A7.5, the estimated parameters $\hat{\theta}_a^{(l)}$ in (7.14) and $\hat{b}_l$ in (7.18) are consistent, i.e., $\hat{\theta}_a^{(l)} \to \theta^{(l)}$ and $\hat{b}_l \to b_l$, as $K \to \infty$.*

Proof of Theorem 7.1: Based on Lemma 7.1, the consistency of $\hat{\theta}_a^{(l)}$ can be proved analogously to the counterpart proof of Theorem 1 in [131]. We only provide the proof for the consistency of $\hat{b}_l$. Under Assumptions A7.1-A7.5, (7.17) holds once n_a, n_b and τ are

obtained as shown in Section 7.3.4. Assuming the noise $\Delta_e(t)$ available and applying the ordinary LSM to (7.17) yields

$$b_l = \left[\frac{1}{K}\sum_{k=0}^{K-1}(\Delta_w(kp+1)-\Delta_e(kp+1))^2\right]^{-1} \cdot \frac{1}{K}\sum_{k=0}^{K-1}(\Delta_w(kp+1)-\Delta_e(kp+1))(\Delta_w(kp+l)-\Delta_e(kp+l)). \quad (7.26)$$

The condition that

$$\frac{1}{K}\sum_{k=0}^{K-1}(\Delta_w(kp+1)-\Delta_e(kp+1))^2 \neq 0$$

is always fulfilled under Assumption A7.2. As $K \to \infty$, $\hat{A}(q)$ converges into $A(q)$ so that $w(t) \to A(q)y(t) = B(q)x(t) + e(t)$. The time-domain expression of (7.11) is

$$\Delta_w(kp+l) = b_l\Delta_x(kp-p) + \Delta_e(kp+l), l = 1,2,\cdots,n_b, \ \forall k \in \mathbb{Z}_+. \quad (7.27)$$

Since $x(t)$ is only possibly correlated with $e(t-d)$ for $d \geq 1$, $\Delta_x(kp-p)$ and $\Delta_e(kp+l)$ for $l = 1,2,\cdots,n_b$ are mutually independent, which, together with (7.27), implies

$$\begin{aligned} E\{\Delta_w(kp+1)\Delta_e(kp+1)\} &= 2\sigma^2, \\ E\{\Delta_w(kp+1)\Delta_e(kp+l)\} &= -\sigma^2\delta(l-2), \\ E\{\Delta_e(kp+1)\Delta_w(kp+l)\} &= -\sigma^2\delta(l-2), \\ E\{\Delta_w(kp+1)\Delta_e(kp+l)\} &= -\sigma^2\delta(l-2). \end{aligned}$$

Therefore, as $K \to \infty$, (7.26) becomes

$$\begin{aligned} b_l &= \left[E\left\{(\Delta_w(kp+1)-\Delta_e(kp+1))^2\right\}\right]^{-1} \\ &\quad \cdot E\{(\Delta_w(kp+1)-\Delta_e(kp+1))(\Delta_w(kp+l)-\Delta_e(kp+l))\} \\ &= \left[E\{\Delta_w^2(kp+1)\} - 2\sigma^2\right]^{-1}\left[E\{\Delta_w(kp+1)\Delta_w(kp+l)\} + \sigma^2\delta(l-2)\right] \quad (7.28) \end{aligned}$$

Comparing (7.18) with (7.28), we have the consistency of $\hat{b}_l$, i.e., $\lim_{K\to\infty}\hat{b}_l = b_l$. □

Theorem 7.2 *Under Assumption A7.6, the estimated inner signal $\hat{X}(n)$ is uniquely determined in (7.25) for a given realization.*

Proof of Theorem 7.2: The uniqueness of $\hat{X}(n)$ in (7.25) requires that the $p(K-1) \times K$ matrix $\mathbf{H}$ has full-column rank, which is true if and only if all the channels $H_i(q)$ for $i = 1,2,\cdots,p$ do not share any common zero except at infinity (Lemma 2 in [71] and Corollary 3.1 in [12]). Specifically, in this context,

$$H_i(q) = h_i(0) + h_i(1)q^{-1}.$$

To make all the channels $H_i(q)$ for $i = 1, 2, \cdots, p$ do not share any common zero except at infinity is equivalent to the condition that the matrix

$$\begin{bmatrix} h_1(0) & h_1(1) \\ h_2(0) & h_2(1) \\ \vdots & \vdots \\ h_p(0) & h_p(1) \end{bmatrix} = \begin{bmatrix} b_1 & \sum_{j=2}^{n_b} b_j \\ b_1 + b_2 & \sum_{j=3}^{n_b} b_j \\ \vdots & \vdots \\ \sum_{j=1}^{n_b-1} b_j & b_{n_b} \\ \sum_{j=1}^{n_b} b_j & 0 \\ \vdots & \vdots \\ \sum_{j=1}^{n_b} b_j & 0 \end{bmatrix}$$

has a trivial null space, which is satisfied by Assumption A7.6, i.e., $\sum_{j=1}^{n_b} b_j \neq 0$. □

7.6 Algorithm and Simulation

This section summarizes the detailed steps of the proposed blind approach and presents a simulated numerical example to illustrate them.

Algorithm:

1. The order n_a and time delay τ are obtained as discussed in Section 7.3.4 by looking at $V^{(p)}(\hat{n}_a)$'s in (7.19) with $l = p$ for different combinations of $\hat{n}_a$ and $\hat{\tau}$.

2. The output $y(t)$ is shifted properly according to $\hat{\tau}$ to make (7.3) hold. The denominator parameters in θ_a are estimated with $n_a = \hat{n}_a$ and $l = p$ from the bias-compensated LSM in (7.14).

3. The order n_b is obtained on the basis of the filtered output $w(t) = \hat{A}(q)y(t)$, as discussed in Section 7.3.4.

4. The numerator parameters $b_2, b_3, \cdots, b_{\hat{n}_b}$ are estimated by the bias-compensated LSM in (7.18).

5. From $\hat{a}_i$ and $\hat{b}_j$, the unmeasurable inner signal $X(n)$ is estimated in (7.25). Owing to the property in (7.1), the fast-rate inner signal $x(t)$ is also available from $\hat{X}(n)$ by piece-wise constant interpolation.

6. The nonlinearity f can be seen from a graph of $\hat{x}(t)$ v.s. $u(t)$. If $f(\cdot)$ has a parametric model, its parameters can be estimated via least-squares nonlinear curve fitting based on the input $U(n)$ and $\hat{X}(n)$, or $u(t)$ and $\hat{x}(t)$, e.g., by the 'lsqcurvefit' function in Matlab Optimization Toolbox.

7. To compensate the errors contained in $\hat{x}(t)$, $G(q)$ may be identified again from $\hat{x}(t)$ and $y(t)$ to compensate the errors contained in $\hat{x}(t)$, e.g., by the 'arx' function in Matlab System Identification Toolbox.

Example 7.1 In Figure 7.2, the process $G_c(s)$ is the same as that in [14] except that it has an additional time delay 0.36 sec and works in a feedback loop with a pure gain controller, i.e.,

$$G_c(s) = \frac{0.4095s + 1.0921}{s^2 + 0.32s + 0.02} e^{-0.36s}, \quad C(z) = 0.1.$$

Here the updating period of the ZOH is $T_h = 0.6$ sec. The nonlinearity f_c is a backlash with deadband 0.1. The process noise is generated by passing zero-mean white noise having variance σ^2 through $1/\left(s^2 + 0.32s + 0.02\right)$. The upper bound of the order n_b is known *a prior* as $n_b^0 = 4$. The fast-sampling ratio is chosen as $p = n_b^0 + 1 = 5$. Thus, the fast-rate process $G(q)$ at the sampling period $h = T/p = 0.12$ sec is

$$G(q) = 0.05597q^{-3} \frac{q^{-1} - 0.7244q^{-2}}{1 - 1.962q^{-1} + 0.9623q^{-2}}.$$

The reference signal r is a random binary sequence with frequency band $[0, 0.5]$ and values ± 1. The simulation duration is 500 sec and the slow-rate inner signal $X(n)$ has around 800 data points to be estimated.

Now we illustrate the estimation of $n_a = 2$, $n_b = 2$ and $\tau = 3$. First, n_a and τ are obtained together by looking at the AIC $V^{(p)}(\hat{n}_a)$ defined in (7.19) for different combinations of $\hat{n}_a$ and $\hat{\tau}$. One typical realization ($\sigma^2 = 0.001$) gives the AICs ($\times 10^{-7}$) in Table 7.1. By

$V^{(p)}(\hat{n}_a)$	$\hat{\tau} = 0$	$\hat{\tau} = 1$	$\hat{\tau} = 2$	$\hat{\tau} = 3$	$\hat{\tau} = 4$
$\hat{n}_a = 1$	12.010	11.235	9.941	8.948	82.172
$\hat{n}_a = 2$	4.619	**0.992**	**0.982**	**1.073**	152.020
$\hat{n}_a = 3$	1.543	1.012	1.735	2.093	152.370
$\hat{n}_a = 4$	1.256	1.111	1.620	1.181	151.480
$\hat{n}_a = 5$	1.148	1.172	1.304	1.204	151.800

Table 7.1: AIC for different combinations of $\hat{n}_a$ and $\hat{\tau}$

looking at the columns with $\hat{\tau} = 1, 2, 3$ in Table 7.1, the AICs do not have significant improvement after $\hat{n}_a = 2$; by looking at the rows with $\hat{n}_a = 2$, the AICs are almost the same for the consecutive shifts $1, 2, 3$ and are smaller than the rest two, i.e., $\hat{\tau} = 3$. The parsimony principle rules out another possible pair $(\hat{n}_a = 3, \hat{\tau} = 3)$. In fact, the pair $(\hat{n}_a = 3, \hat{\tau} = 3)$ yields the almost same inner signal estimation as our choice $(\hat{n}_a = 2, \hat{\tau} = 3)$. Second, after shifting $y(t)$ by $\hat{\tau}$ and estimating the parameters in $A(q)$, the filtered output

$w(t) = \hat{A}(q)y(t)$ is formed; then, the true order n_b is obtained, i.e., $\hat{n}_b = 2$, because $p = 5$ and the variances of $\Delta_w^{(l)}$'s have three almost same smallest numbers ($\times 10^{-7}$) in Table 7.2.

l	1	2	3	4	5
$Var\left\{\Delta_w^{(l)}\right\}$	31.591	59.353	0.9974	0.9920	0.9463

Table 7.2: The variance of $\Delta_w^{(l)}$

Remark: Some modifications are perhaps necessary in estimating τ. For instance, if $\tau = 1$, one realization ($\sigma^2 = 0.001$) gives the AICs ($\times 10^{-7}$) in Table 7.3. The smallest $V^{(p)}(\hat{n}_a)$'s in Table 7.3 are at three inconsecutive shifts 0, 1, and 4; thus, the shift 4 needs to be regarded as -1, because the sole information of $y(t)$ cannot tell the difference between τ and $(\tau + kp)$ for $k \in \mathbb{Z}_+$.

$V^{(p)}(\hat{n}_a)$	$\hat{\tau} = 0$	$\hat{\tau} = 1$	$\hat{\tau} = 2$	$\hat{\tau} = 3$	$\hat{\tau} = 4$
$\hat{n}_a = 2$	0.9948	1.0182	122.03	3.2380	1.0146

Table 7.3: AICs for time delay estimation

$\sigma^2 \times 10^{-2}$	SNR	$a_1 = -1.9620$	$a_2 = 0.9623$	$b_2 = 0.7244$
0	∞	-1.9620	0.9623	0.7244
0.01	16.2627	-1.9623 ± 0.0018	0.9626 ± 0.0019	0.7240 ± 0.0027
0.05	7.2287	-1.9634 ± 0.0038	0.9638 ± 0.0038	0.7242 ± 0.0053
0.1	5.1003	-1.9651 ± 0.0054	0.9656 ± 0.0055	0.7247 ± 0.0085
0.5	2.4956	-1.9723 ± 0.0112	0.9735 ± 0.0113	0.7252 ± 0.0240
1	1.8889	-1.9781 ± 0.0150	0.9799 ± 0.0151	0.7232 ± 0.0348

Table 7.4: Estimated parameters and their standard deviations

Next, we investigate the performance of the parameter and inner signal estimation by multiple Monte Carlo simulations. Table 7.4 presents the averaged estimates of the parameters a_1, a_2 and b_2 and their standard deviations for different noise levels; 100 Monte Carlo simulations are performed for each non-zero noise level. Here the signal-to-noise ratio is defined as SNR $= \|y_0(t)\|_2 / \|v(t)\|_2$, where $y_0(t)$ is the noise-free component of $y(t)$ and $\|\cdot\|_2$ denotes the Euclidean norm. As expected, the estimates are consistent. Using the estimated parameters, the unmeasurable inner signal $X(n)$ is obtained in (7.25); a graph of the controller output $u(n)$ v.s. $\hat{X}(n)$ from one typical realization with $\sigma^2 = 0.001$ is shown in Figure 7.4. As a comparison, Figure 7.5 shows the graph of $u(n)$ v.s. the true inner signal $X(n)$. Since the gain of $G(q)$ is impossible to get in practice from the information of the

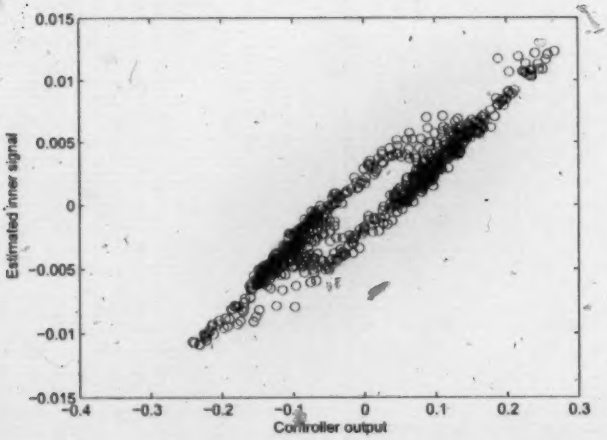

Figure 7.4: A graph of the controller output $u(n)$ v.s. the estimated inner signal $\hat{X}(n)$

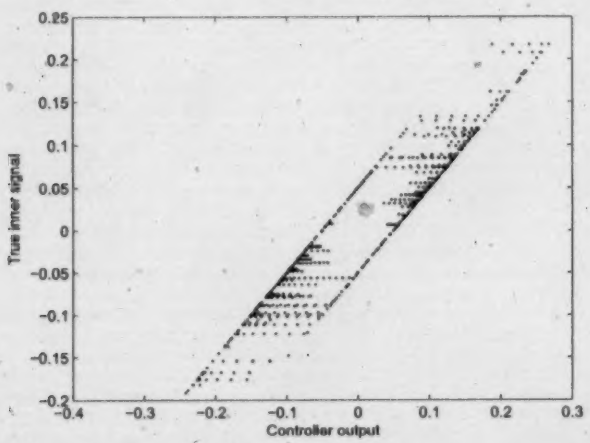

Figure 7.5: A graph of the controller output $u(n)$ v.s. the true inner signal $X(n)$

output only, the vertical axes of Figures 7.4 and 7.5 have different scales. Nevertheless, the nonlinearities in Figures 7.4 and 7.5 have a good match in terms of the backlash structure; the deadband is read from the graph to be approximately 0.1.

For the purpose of comparison, the gain of $G(q)$ is assumed to be known, and $\hat{X}(n)$ is scaled properly. The error between the two signals is measured numerically by a fitness in

$\sigma^2 \times 10^{-2}$	SNR	F_{LSM}	F^0_{LSM}	F_{INV}	F^0_{INV}
0	∞	100	100	100	100
0.01	16.2627	95.1989 ± 1.6295	97.2922	93.8494 ± 1.4933	95.6009
0.05	7.2287	89.4944 ± 3.8574	93.8981	86.4244 ± 3.5078	90.1067
0.1	5.1003	85.0520 ± 5.2490	91.3634	80.7704 ± 4.7868	86.0497
0.5	2.4956	62.1737 ± 16.5220	80.9620	53.0911 ± 15.5598	69.2815
1	1.8889	49.0576 ± 24.8790	73.2331	36.0619 ± 24.0173	57.4234

Table 7.5: Averaged fitnesses and their standard deviations: the standard deviations of F^0_{LSM} and F^0_{INV} are omitted

% (see 'compare' command in Matlab System Identification Toolbox),

$$F(\hat{X}_s, X) = 100\left(1 - \frac{\left\|\hat{X}_s(n) - X(n)\right\|_2}{\|X(n) - E\{X(n)\}\|_2}\right). \tag{7.29}$$

Here $\hat{X}_s(n)$ stands for the estimated inner signal after scaling. Table 7.5 lists the averaged fitnesses and their standard deviations from the same Monte Carlo simulations as those in Table 7.4. F_{LSM} is the fitness between $X(n)$ and $\hat{X}_s(n)$ obtained from the LSM in (7.25). The upper bound of F_{LSM}, denoted by F^0_{LSM}, is calculated by using the true parameters a_i and b_j in (7.25); the standard deviation of F^0_{LSM} is relatively small and is omitted here. In this example, $G(q)$ is minimum phase so that $X(n)$ can also estimated by passing $y(t)$ through the direct inverse of $\hat{G}(q)$ and downsampling the resulted signal $\hat{x}(t)$ by $p = 5$. The corresponding fitness and its upper bound are denoted by F_{INV} and F^0_{INV}, respectively. Eq. (7.25) reduces the noise effect in the inner signal estimation, as F^0_{LSM} and F_{LSM} are always larger than F^0_{INV} and F_{INV}, respectively. The difference between F^0_{LSM} and F_{LSM} is getting larger as the SNR decreases; this is due to the propagation of the errors in parameter estimation into the inner signal estimation. □

7.7 MR Damper Modeling

The proposed blind approach is applied for modeling of MR dampers. First, we briefly introduce the Hammerstein model of MR dampers. Next, the setup and result of the modeling experiment are presented.

7.7.1 Hammerstein Model

The nonparametric model for MR dampers proposed in [124] has demonstrated two merits so far: (i) the model can be numerically solved much faster than the existing parametric

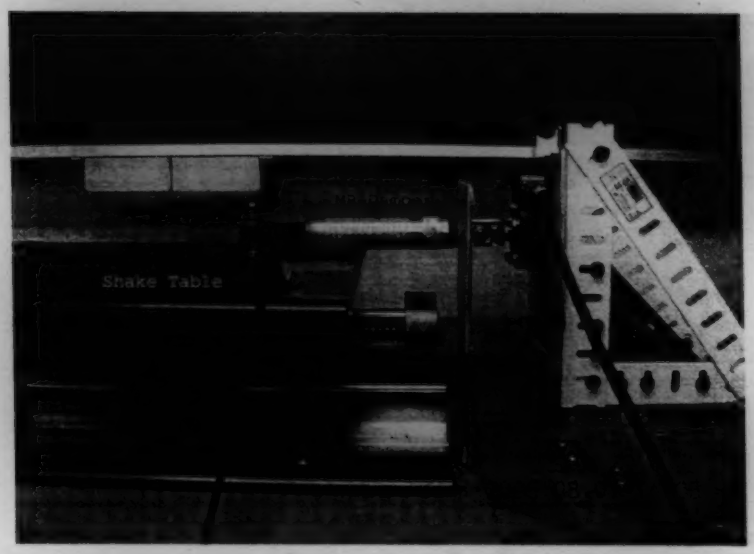

Figure 7.6: Experimental devices

models; (ii) the stability of an MR damper control system can be proved by adopting the nonparametric model [125]. If currents/voltages of MR dampers are constants, the nonparametric model becomes a Hammerstein system depicted in Figure 7.1. Here the input $u(t)$ and output $y(t)$ stand for the velocity and damping force, respectively. Song, Ahmadian & Southward [124] proposed a first-order model for the linear system,

$$G(q) = \frac{b_1 q^{-1}}{1 + a_1 q^{-1}}, \tag{7.30}$$

and three candidate functions for the nonlinearity,

$$f(u) = c_1 \tanh(c_0 u), \tag{7.31}$$

$$f(u) = c_1 \mathrm{sgn}(u) \left[1 - \exp\left(-c_0 |u|\right)\right],$$

$$f(u) = \frac{\left(c_0 + c_1 |u - c_3|\right)^{c_2(u-c_3)} - \left(c_0 + c_1 |u - c_3|\right)^{-c_2(u-c_3)}}{c_0^{c_2(u-c_3)} + c_0^{-c_2(u-c_3)}}.$$

Our objective is to design an identification experiment and estimate $G(q)$ and $f(\cdot)$ from the measured damping force $y(t)$ and velocity $u(t)$ by the proposed blind approach.

7.7.2 Experiment Setup

Experimental devices and a diagram of the experimental setup are depicted in Figures 7.6 and 7.7, respectively. Two ends of the MR damper (RD-1097-01) provided by Lord Corp. are

connected to the shake table and ground, respectively. The shake table generates necessary vibrations; in other words, the velocity of the MR damper is determined by the displacement of the shake table. Since the shake table weights about 60 lbs leading to a large inertia, it has to be controlled under a closed-loop operation. The proportional-derivative (PD) controller in Figure 7.7 is implemented in Computer #1, and reads the displacement by countering turns of a circulating shaft and sends out currents to drive the shake table at sampling period 0.001 sec. Simultaneously, Computer #2 reads the damping force via a strain meter and the displacement via an infrared sensor at sampling period 0.005 sec. After downsampling the measurements from Computer #1 by a factor 5, we synchronize all measurements from the two computers by comparing the two displacement measurements. Eventually, displacement measurements from Computer #2 are discarded because they are much nosier. No velocity sensor is available so that the velocity is calculated as the first derivative of displacement measurements from Computer #1. The voltage of the MR damper is fixed to 1.25 v.

Assumption A7.1 in Section 7.2 requires the velocity to be piece-wise constant for p consecutive samples. We let the desired displacement in Figure 7.7 take uniformly-distributed random values within the range [-1.5, 1.5] cm and have a constant increment every 0.2 sec. As a result, the velocity is approximately piece-wise constant for every 40 samples (the sampling period h is 0.005 sec). Figure 7.8 shows some enlarged parts of the measured displacement, the calculated velocity and the measured damping force (bottom to top). The duration 0.2 sec is confined by the closed-loop settling time, as the shake table has a rather large inertia.

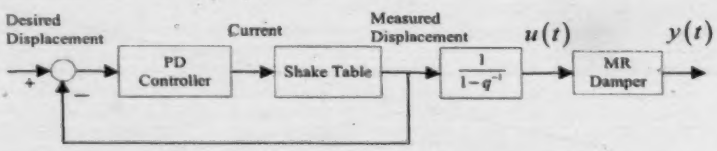

Figure 7.7: A diagram of the experimental setup

7.7.3 Experiment Result

The selected experimental data with 3500 samples are presented in Figure 7.9. The first half of the data is used for parameter estimation, while the other half is for model validation.

Let us first look at some assumptions. By our design in Section 7.7.2, Assumption A7.1 is approximately satisfied, i.e., the velocity is approximately piece-wise constant for every 40

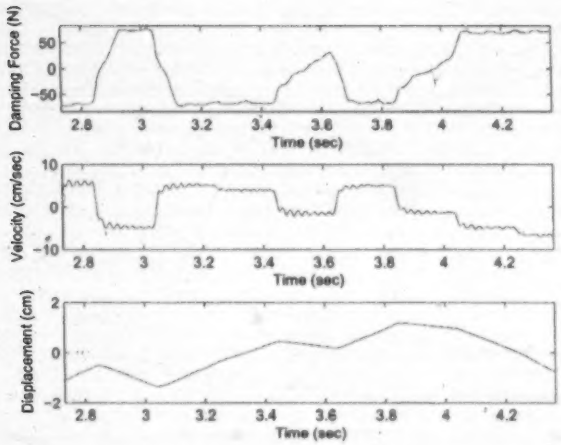

Figure 7.8: Some enlarged parts of experimental data

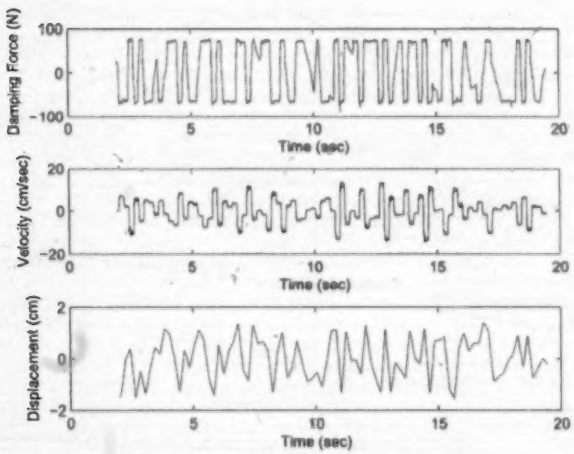

Figure 7.9: Selected experimental data

samples. The integer p could be as large as 40; however, a larger p implies that fewer data points are exploited and the inner signal estimation needs higher computational costs. The linear system $G(q)$ in (7.30) is expected to be the first order one, i.e., $n_a = 1$ and $n_b = 1$. Hence, $p = 5$ (a factor of 40) seems a well-balanced choice to safely satisfy Assumption A7.5 and meet the above consideration on the data length and computational cost.

The identification algorithm in Section 7.6 is proceeded as follows. First, the order n_a and time delay τ are determined from Table 7.6 that lists $V^{(p)}(\hat{n}_a)$'s in (7.19) under different combinations of $\hat{n}_a$ and $\hat{\tau}$. Looking at the rows of Table 7.6, $V^{(p)}(\hat{n}_a)$ for $\hat{\tau} = 4$

$V^{(p)}(\hat{n}_a)$	$\hat{\tau}=0$	$\hat{\tau}=1$	$\hat{\tau}=2$	$\hat{\tau}=3$	$\hat{\tau}=4$
$\hat{n}_a=1$	7.927	11.113	12.713	10.844	5.027
$\hat{n}_a=2$	13.866	18.366	21.837	15.012	5.171
$\hat{n}_a=3$	13.391	18.420	21.618	15.027	5.063
$\hat{n}_a=4$	13.379	18.483	21.711	15.378	5.092
$\hat{n}_a=5$	12.970	18.558	21.824	15.406	5.121

Table 7.6: The AICs under different combinations of $\hat{n}_a$ and $\hat{\tau}$

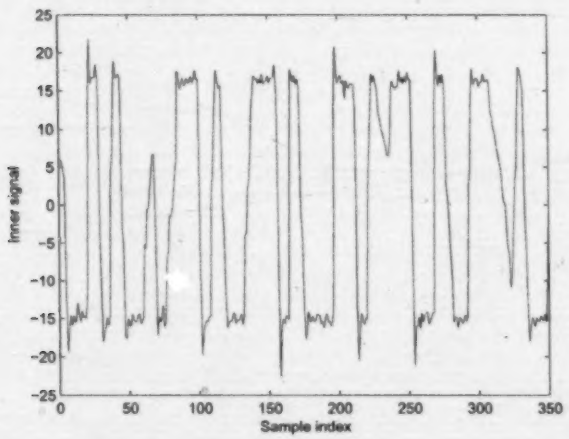

Figure 7.10: The estimated inner signal $X(n)$

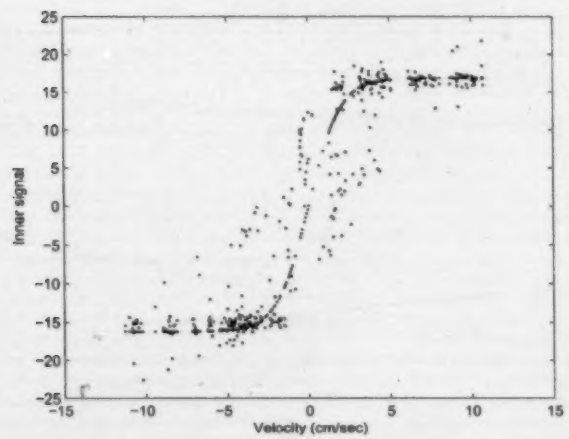

Figure 7.11: The nonlinearity (dots) revealed by $U(n)$ and $\hat{X}(n)$, and the estimated nonlinearity (smooth line)

achieves the smallest number at each row. In the column of $\hat{\tau} = 4$, $V^{(p)}(\hat{n}_a)$ is almost the same for all $\hat{n}_a$'s. In fact, $\hat{n}_a = 1$ and $\hat{\tau} = 4$ lead to the smallest AIC. With $\hat{n}_a = 1$, $\hat{\tau} = 4$ and $l = p + \hat{\tau}$, the denominator parameter is obtained in (7.14):

$$\hat{A}(q) = 1 - 0.8433q^{-1}. \tag{7.32}$$

There is no need to estimate the numerator parameters, because $\hat{n}_b = 1$ and $b_1 = 1$ (see Section 7.2). Second, the slow-rate inner signal $X(n)$ is estimated in (7.25); $\hat{X}(n)$ is shown in Figure 7.10. Third, the shape of the nonlinearity is revealed from $U(n)$ and $\hat{X}(n)$ displayed by dots in Figure 7.11. It seems that $f(\cdot)$ in (7.31) would be sufficient to capture the revealed nonlinearity,

$$\hat{f}(U) = 0.5080 \tanh(16.2320U). \tag{7.33}$$

Interpolating $\hat{X}(n)$ by the property in (7.1) gives $\hat{x}(t)$, the estimate of the inner signal $x(t)$ with sampling period 0.005 sec. The linear system $G(q)$ is ready to be identified from $\hat{x}(t)$ and $y(t)$,

$$\hat{G}(q) = \frac{0.5357q^{-1}}{1 - 0.8856q^{-1}}. \tag{7.34}$$

Finally, the simulated damping force $\hat{y}(t)$ is obtained by passing $u(t)$ through $\hat{f}(\cdot)$ and $\hat{G}(q)$, i.e.,

$$\hat{y}(t) = \hat{G}(q)\hat{f}(u(t)).$$

Figure 7.12 compares $\hat{y}(t)$ with the measured damping force $y(t)$. The fitness between $y(t)$ and $\hat{y}(t)$, $F(\hat{y}, y)$ defined in (7.29), is 70.5102%, The other half of the data in Figure 7.9 is used for the cross validation. The corresponding $y(t)$ and $\hat{y}(t)$ are compared in Figure 7.13 with fitness 63.1926%.

Since the actual inner signal $x(t)$ or $X(n)$ is unavailable in the experiment, the fitness between the measured and simulated damping forces is the single index to evaluate the model quality. In Figures 7.12 and 7.13, the estimated Hammerstein model consisting of $\hat{f}(\cdot)$ in (7.33) and $\hat{G}(q)$ in (7.34) performs very well in terms of dynamics tracking, but has relatively large errors at some smaller peaks. Numerically, the Hammerstein model is validated by the fair fitness 63.1926% between the simulated and measured outputs in the cross validation. The modeling errors may arise from the noises in the measured damping force and the approximation in achieving the piece-wise constant velocity.

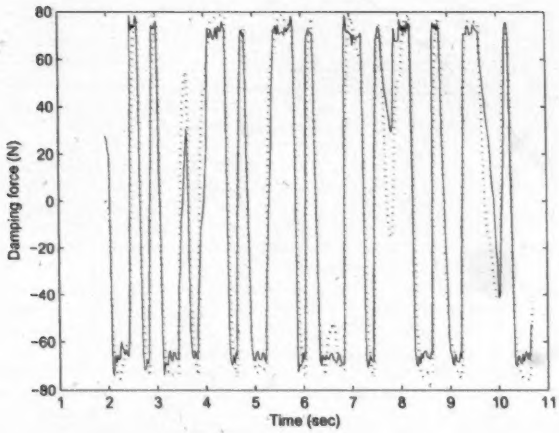

Figure 7.12: The measured (solid) and simulated (dotted) damping forces using the estimation data: fitness = 70.5102%

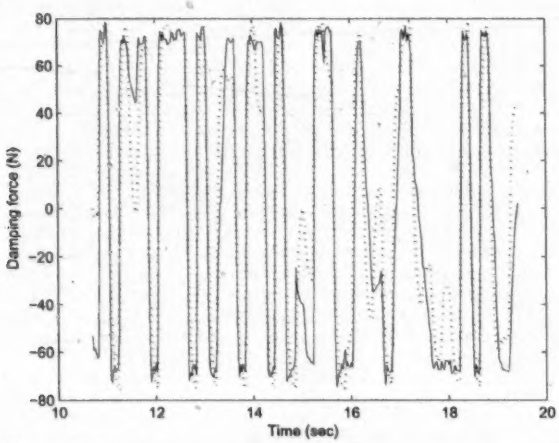

Figure 7.13: The measured (solid) and simulated (dotted) damping forces using the validation data: fitness = 63.1926%

7.8 Conclusion

We have proposed a new blind approach to identification of Hammerstein systems. Inputs of Hammerstein systems are assumed to be piece-wise constant for certain consecutive samples. Such an assumption can be satisfied by user's design, e.g., in the MR modeling experiment in Section 7.7, or by output fast-sampling, e.g., in Example 7.1. In a laboratory experiment, the proposed approach has been applied to identify a Hammerstein model for MR dampers.

An interesting issue for further study is to compare the proposed approach with the existing ones [131, 14] by simulation comparison and theoretical analysis. The two differences with the existing approaches stated in Section 7.1 imply that the proposed approach may have a better performance in reducing noise effects.

Chapter 8

Conclusion

The thesis consists of two parts, namely, cyclo-stationary signal analysis and its applications in system identification. To get a clear overall picture, we list the main contributions of the thesis as follows.

Main contributions:

1. Cyclo-period estimation (Chapter 3): A new method, named as the variability method, is proposed to estimate the cyclo-period of a discrete-time cyclo-stationary signal. The variability method has many attractive properties, e.g., it is not sensitive to stationary noises. These properties are analyzed and compared with three existing cyclo-period estimation methods via simulation and real-life examples.

2. Cyclo-statistic estimation (Chapter 4): The first- and second-order cyclo-statistic estimators are summarized, namely, the estimators of the time-varying mean/correlation and cyclic correlation/spectrum. A new cyclic spectrum estimator, the blocking-based estimator, is proposed. The rationale of an implementation shortcut for the cyclic mean/correlation/spectrum estimator is explored from the relationship between cyclo-stationarity and quasi-stationarity. Performance of the cyclo-statistic estimators is validated via simulation examples.

3. Cyclo-spectral theory (Chapter 5): Two problems are studied for the spectral theory of discrete-time cyclo-stationary signals: the cyclospectrum representation and the cyclospectrum transformation by linear multirate systems. The first contribution, which is also of some tutorial value, is to summarize four types of cyclospectra and find their interrelationships. In the literature, the problem of cyclospectrum transformation by linear systems was investigated only for some specific configurations and was usually developed with inordinate complexities due to lack of a systematic approach. The

second contribution is to attack the problem of the cyclospectrum transformation in the framework of multirate systems using the blocking technique.

4. FIR modeling for EIV/closed-loop systems (Chapter 6): A non-parametric approach, the CCRA, yields asymptotically unbiased and consistent FIR models for EIV and closed-loop systems. A complete study of the CCRA is developed, including the statistical performance of the estimated FIR model. Frequency-domain expressions of the statistical performance provide guidelines in designing a class of cyclo-stationary signals for modeling. Effectiveness and properties of the CCRA are validated and illustrated by numerical examples.

5. Blind identification of Hammerstein systems (Chapter 7): The first contribution is to propose a new blind approach to identification of Hammerstein systems. The new approach has two main differences with the existing blind approaches. (i) The noise-corrupted cases are considered instead of the noise-free ones, leading to a new series of static EIV systems. (ii) The inner signal is estimated by a least-squares method borrowed from the blind equalization; by doing so, the noise effect is reduced. The second contribution lies at modeling of MR dampers. We design a real-time identification experiment for MR dampers and build a Hammerstein model by the proposed blind approach.

Bibliography

[1] K. Abed-Meraim, W. Qiu and Y. Hua, Blind system identification, *Proc. of the IEEE*, **85(8)**, 1310-1322, 1997.

[2] S. Akkarakaran and P. P. Vaidyanathan, Bifrequency and bispectrum maps: a new look at multirate systems with stochastic inputs, *IEEE Trans. Signal Processing*, **48(3)**, 723-736, 2000.

[3] V. G. Alekseev, Estimating the spectral densities of a Gaussian periodically correlated stochastic process, *Problems of Information Transmission*, **24(2)**, 109-115, 1988.

[4] K.J. Åström and T. Hägglund, Automatic tuning of simple regulators with specifications on phase and amplitude margins, *Automatica*, **20(5)**, 645-651, 1984.

[5] B. D. O. Anderson and M. Deistler, Identifiability of dynamic errors-in-variables models, *J. Times Ser. Anal.*, **5(1)**, 1-13, 1984.

[6] J. M. M. Anderson and G. B. Giannakis, Noisy input/output system identification using cumulants and the Steiglitz-McBride algorithm, *IEEE Trans. Signal Processing*, **44(2)**, 1021-1024, 1996.

[7] J. Antoni, P. Wagstaff and J. C. Henrio, H_α – A consistent estimator for frequency response functions with input and output noise, *IEEE Trans. Instrum. Meas.*, **53(2)**, 457-465, 2004.

[8] J. Antoni, F. Bonnardot, A. Raad and M. El Badaoui, Cyclostationary modelling of rotating machine vibration signals, *Mechanical Systems and Signal Processing*, **18**, 1285-1314, 2004.

[9] E. W. Bai, An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems, *Automatica*, **34(3)**, 333-338, 1998.

[10] E.W. Bai, Identification of linear systems with hard input nonlinearities of known structure, *Automatica*, **38(5)**, 853-860, 2002.

[11] E.W. Bai, Decoupling the linear and nonlinear parts in Hammerstein model identification, *Automatica*, **40(4)**, 671-676, 2004.

[12] E.W. Bai and Z. Ding, Invertibility of sampled data systems, *IEEE Trans. Circuits Systems I*, **47(3)**, 279-289, 2000.

[13] E.W. Bai and M. Fu, Blind system identification and channel equalization of IIR systems without statistical information, *IEEE Trans. Signal Processing*, **47(7)**, 1910-1921, 1999.

[14] E.W. Bai and M. Fu, A blind approach to Hammerstein model identification, *IEEE Trans. Signal Processing*, **50(7)**, 1610-1619, 2002.

[15] E.W. Bai, Q. Li and S. Dasgupta, Blind identifiability of IIR systems, *Automatica*, **38(1)**, 181-184, 2002.

[16] A. Balestrino, A. Landi, M. Ould-Zmirli and L. Sani, Automatic nonlinear autotuning method for Hammerstein modeling of electrical drives, *IEEE Trans. Industrial Electronics*, **48(3)**, 645-655, 2001.

[17] W.L. Bialkowski, Dreams vs. reality: A view from both sides of the gap, in *Proc. of Control Systems*, 283-294, Whistler, BC, Canada, 1992.

[18] S. A. Billings and S. Y. Fakhouri, Non-linear system identification using the Hammerstein model, *Int. J. Syst. Sci.*, **10(5)**, 567-578, 1979.

[19] S. Björklund, *A Survey and Comparison of Time-Delay Estimation Methods in Linear Systems*, Ph.D. Thesis, Linköping, Sweden: Linköping University, 2003.

[20] P. Bloomfield, H. L. Hurd and R. B. Lund, Periodic correlation in stratospheric ozone data, *J. Times Ser. Anal.*, **15(2)**, 127-150, 1994.

[21] B. Boashash, Time-frequency signal analysis: past, present and future trends, in *Control and Dynamic Systems*, C.T. Leondes, Ed., Academic Press, **78**, 1-69, 1996.

[22] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, Englewood Cliffs, NJ: Prentice-Hall, 1994.

[23] G. E. P. Box and J. F. MacGregor, The analysis of closed-loop dynamic-stochastic systems, *Technometrics*, **16**, 391-398, 1974.

[24] R. A. Boyles and W. A. Gardner, Cycloergodic properties of discrete-parameter nonstationary stochastic processes, *IEEE Trans. Inform. Theory*, **29(1)**, 105-114, 1983.

[25] W. A. Brown and H. H. Loomis, Digital implementations of spectral correlation analyzers, *IEEE Trans. Signal Processing*, **41(2)**, 703-720, 1993.

[26] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, Upper Saddle River, NJ: Prentice-Hall, 1998.

[27] F. Chang and R. Luus, A noniterative method for identification using Hammerstein model, *IEEE Trans. Automat. Contr.*, **16(5)**, 464-468, 1971.

[28] T. Chen and B. A. Francis, *Optimal Sampled-data Control Systems*, London: Springer, 1995.

[29] T. Chen, L. Qiu and E. W. Bai, General multirate building structures with application to nonuniform filter banks, *IEEE Trans. Circuits and Syst. II*, **45(8)**, 948-958, 1998.

[30] T. Chen, Nonuniform multirate filter bank: analysis and design with an $\mathcal{H}_\infty$ performance measure, *IEEE Trans. Signal Processing*, **45(3)**, 572–582, 1997.

[31] S. W. Cho, H. J. Jung, J. H. Lee and I. W. Lee, Smart passive system based on MR damper. *JSSI 10th Anniversary Symp. on Performance of Response Controlled Buildings*, Yokohama, Japan, November 2004.

[32] S. B. Choi and S. K. Lee, A hysteresis model for the field-dependent damping force of a magnetorheological damper. *J. Sound and Vibration*, **245(2)**, 375-383, 2001.

[33] M. A. A. S. Choudhury, N. F. Thornhill and S. L. Shah, Modeling valve stiction, *Control Engineering Prac.*, **13(5)**, 641-658, 2005.

[34] J. W. Cooley and J. W. Tukey, An algorithm for the machine computation of complex Fourier series, *Math. Computation*, **9**, 297-301, 1965.

[35] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Engelwood Cliffs, NJ: Prentice-Hall, 1983.

[36] A. V. Dandawaté and G. B. Giannakis, Nonparametric polyspectral estimators for kth-order (almost) cyclostationary processes, *IEEE Trans. Inform. Theory*, **40(1)**, 67-84, 1994.

[37] A. V. Dandawaté and G. B. Giannakis, Statistical tests for presence of cyclostationarity, *IEEE Trans. Signal Processing*, **42(9)**, 2355-2369, 1994.

[38] A. V. Dandawaté and G. B. Giannakis, Asymptotic theory of mixed time averages and kth-order cyclic-moment and cumulant statistics, *IEEE Trans. Inform. Theory*, **41(1)**, 216-232, 1995.

[39] C.E. Davila, An efficient recursive total least squares algorithm for FIR adaptive filtering, *IEEE Trans. Signal Processing*, **42(2)**, 268-280, 1994.

[40] B. S. Dayal and J. F. MacGregor, Multi-output process identification, *J. Proc. Cont.*, **7(4)**, 269-282, 1997.

[41] D. Dehay and H. L. Hurd, Representation and estimation for periodically and almost periodically correlated random processes, in *Cyclostationarity in Communications and Signal Processing*, W. A. Gardner, Ed., Piscataway, NJ: IEEE Press, 1994.

[42] A. Delopoulos and G. B. Giannakis, Consistent identification of stochastic linear systems with noisy input-output data, *Automatica*, **30(8)**, 1271-1294, 1994.

[43] E. J. Dempsey and D. T. Westwick, Identification of Hammerstein models with cubic spline nonlinearities, *IEEE Trans. Biomed. Eng.*, **51(2)**, 237-245, 2004.

[44] L. Desborough and R. Miller, Increasing customer value of industrial control performance monitoring - Honeywell's experience, in *Proc. of Chemical Process Control VI.*, 172-192, Tuscon, AZ, USA, 2001.

[45] P. Erdos, J. Suranyi and B. Guiduli, *Topics in the Theory of Numbers*, New York: Springer Verlag, 2003.

[46] E. Eskinat, S. H. Johnson and W. L. Luyben, Use of Hammerstein models in identification of nonlinear systems, *A.I.Ch.E. J.*, **37(2)**, 255-268, 1991.

[47] F. Flagiello, L. Izzo and A. Napolitano, A computationally efficient and interference tolerant nonparametric algorithm for LTI system identification based on higher order cyclostationarity, *IEEE Trans. Signal Processing*, **48(4)**, 1040-1052, 2000.

[48] N. J. Fliege, *Multirate Digital Signal Processing*, New York: Wiley, 1994.

[49] C. H. O. Fontes and M. Embirucu, Multivariable correlation analysis and its application to an industrial polymerization reactor, *Computers and Chemical Engineering*, **25(2)**, 191-201, 2001.

[50] B. A. Francis and T. T. Georgiou, Stability theory for linear time-invariant plants with periodic digital controllers, *IEEE Trans. Automat. Control*, **33(9)**, 820–832, 1988.

[51] J. E. Freund *Mathematical statistics*, 5th ed. Englewood Cliffs, NJ: Prentice Hall, 1992.

[52] W. A. Gardner, Measurement of spectral correlation, *IEEE Trans. Inform. Theory*, **34(5)**, 1111-1123, 1986.

[53] W. A. Gardner, *Statistical Spectral Analysis: a Nonprobabilistic Theory*, Engelwood Cliffs, NJ: Prentice-Hill, 1988.

[54] W. A. Gardner, *Introduction to Random Processes with Application to Signals and Systems*, 2nd ed., New York: McGraw-Hill, 1990.

[55] W. A. Gardner, Identification of systems with cyclostationary input and correlated input/output measurement noise, *IEEE Trans. Automat. Control*, **35(4)**, 449-452, 1990.

[56] W. A. Gardner, An introduction to cyclostationary signals, in *Cyclostationarity in Communications and Signal Processing*, W. A. Gardner, Ed., Piscataway, NJ: IEEE Press, 1994.

[57] W. A. Gardner, Identification of polyperiodic nonlinear systems, *Signal Processing*, **46(1)**, 75-83, 1995.

[58] W. A. Gardner and T. L. Archer, Exploitation of cyclostatinoarity for identifying the Volterra kernels of nonlinear systems, *IEEE Trans. Inform. Theory*, **39(2)**, 535-542, 1993.

[59] M. J. Genossar, H. Lev-Ari and T. Kailath, Consistent estimation of the cyclic autocorrelation, *IEEE Trans. Signal Processing*, **42(3)**, 595-603, 1994.

[60] G. B. Giannakis, Polyspectral and cyclostationary approaches for identification of closed-loop systems, *IEEE Trans. Automat. Control*, **40(5)**, 882-885, 1995.

[61] G. B. Giannakis, Cyclostationary signal analysis, in *Digital Signal Processing Handbook*, Boca Raton: CRC Press LLC, 1999.

[62] E. G. Gladyshev, Periodically correlated random sequence, *Soviet Math.*, **2**, 385-388, 1961.

[63] J. C. Gomez and E. Baeyens, Identification of block-oriented nonlinear systems using orthonormal bases, *J. Process Control*, **14(6)**, 685-697, 2004.

[64] J. C. Gomez and E. Baeyens, Subspace-based identification algorithms for Hammerstein and Wiener models, *Eur. J. Control*, **11**, 127-136, 2005.

[65] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, Maryland: Johns Hopkins University Press, 1996.

[66] W. Greblicki and M. Pawlak, Identification of discrete Hammerstein systems using kernel regression estimation. *IEEE Trans. Automat. Contr.*, **31(1)**, 74-77, 1986.

[67] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, 2nd ed. Cambridge, England: Cambridge University Press, 1988.

[68] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, New York: Wiley, 1996.

[69] F. Hlawatsch and G. F. Boudreaux Bartels, Linear and Qaudratic Time-Frequency Signal Representations, *IEEE Signal Processing Magazine*, **9**, 21-67, 1992.

[70] L. J. Herbst, The statistical Fourier analysis of variances, *J. R. Statist. Soc.*, **B(27)**, 159-165, 1965.

[71] Y. Hua and M. Wax, Strict identifiability of multiple FIR channels driven by an unknown arbitrary sequence, *IEEE Trans. Signal Processing*, **44(4)**, 756-759, 1996.

[72] H. L. Hurd and N. L. Gerr, Graphical methods for determining the presence of periodic correlation, *J. Times Ser. Anal.*, **12(4)**, 337-350, 1991.

[73] H. L. Hurd, Nonparametric time series analysis for periodically correlated processes, *IEEE Trans. Inform. Theory*, **35(2)**, 350-359, 1989.

[74] A. Janczak, *Identification of Nonlinear Systems Using Neural Networks and Polynomial Models: A Block-Oriented Approach*, New York: Springer-Verlag, 2005.

[75] G. M. Jenkins and D. G. Watts, *Spectral Analysis and its Applications*, San Francisco: Holden-Day, 1968.

[76] H. Kagiwada, L. Sun, A. Sano and W. Liu, Blind identification of IIR model based on output over-sampling, *IEICE Trans. Fundamentals*, **E81-A(11)**, 2350-2360, 1998.

[77] T. Kailath, A. H. Sayed and B. Hassibi, *Linear Estimation*, Upper Saddle River, NJ: Prentice Hall, 2000.

[78] P. P. Khargonekar, K. Poolla and A. Tannenbaum, Robust control of linear time-invariant plants using periodic compensation, *IEEE Trans. Automat. Control*, **30(11)**, 1088-1096, 1985.

[79] B. Lall, S. D. Joshi and R. K. P. Bhatt, Second-order statistical characterization of the filter bank and its elements, *IEEE Trans. Signal Processing*, **47(6)**, 1745-1749, 1999.

[80] M. Lelic and Z. Gajic, A reference guide to PID controllers in the Nineties, *Proc. of IFAC Workshop*, Terrassa, Spain, 2000.

[81] Y. C. Liang, A. R. Leyman and B. H. Soong, (Almost) periodic FIR system identification using third-order cyclic-statistics, *Electronics Letters*, **33(5)**, 356-357, 1997.

[82] W. M. Ling and D. E. Rivera, Multivariable impulse response estimation via correlation analysis and its application to automated system identification, *Proc. of the 11th IFAC Symp. on System Identification*, 1399-1404, Fukuoka, Japan, 1997.

[83] T. Liu and T. Chen, Design of multichannel nonuniform transmultiplexers using general building blocks, *IEEE Trans. Signal Processing*, **49(1)**, 91-99, 2001.

[84] L. Ljung, Some limit results for functionals of stochastic processes, *Report LiTh-ISY-I-0167*, Sweden: Dept. of Electrical Engineering, Linköping University, 1977.

[85] L. Ljung, *System Identification: Theory for the User*, 2nd ed., Englewood Cliffs, NJ: Prentice Hall, 1999.

[86] L. Ljung, *The System Identification Toolbox: The Manual*, 7th ed. Natick, MA: The MathWorks Inc., 2005.

[87] C. M. Loeffler and C. S. Burrus, Optimal design of periodically time-varying and multirate digital filters, *IEEE Trans. Acoust., Speech, Signal Processing*, **32(5)**, 991-997, 1984.

[88] W. L. Luyben and E. Eskinat, Nonlinear auto-tune identification, *Int. J. Control*, **59(3)**, 595-626, 1994.

[89] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*, New York: McGraw-Hill, 2000.

[90] D. E. K. Martin and B. Kedem, Estimation of the period of periodically correlated sequences, *J. Times Ser. Anal.*, **14(2)**, 193-205, 1993.

[91] D. E. K. Martin, Detection of periodic autocorrelation in time series data via zero-crossing, *J. Times Ser. Anal.*, **20(4)**, 435-452, 1999.

[92] D. Mattera and L. Paura, Higher-order cyclostationarity-based methods for identifying Volterra systems by input-output noisy measurements, *Signal Processing*, **67(1)**, 77-98, 1998.

[93] D. Mattera, Identification of polyperiodic Volterra systems by means of input-output noisy measurements, *Signal Processing*, **75(1)**, 41-50, 1999.

[94] A. C. McCormick and A. K. Nandi, Cyclostationarity in rotating machine vibrations, *Mechanical Systems and Signal Processing*, **12(2)**, 225-242, 1998.

[95] J. M. Mendel, Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications, *Proc. of the IEEE*, **79(3)**, 278-305, 1991.

[96] R. A. Meyer and C. S. Burrus, A unified analysis of multirate and periodically time-varying digital filters, *IEEE Trans. Circuits Syst.*, **22(3)**, 162-168, 1975.

[97] D. D. Muresan and T. W. Parks, Orthogonal, exactly periodic subspace decomposition, *IEEE Trans. Signal Processing*, **51(9)**, 2270-2279, 2003.

[98] K. S. Narendra and P. G. Gallman, An iterative method for the identification of nonlinear systems using a Hammerstein model, *IEEE Trans. Automat. Contr.*, **11(3)**, 546-550, 1966.

[99] S. Ohno and H. Sakai, Optimization of filter banks using cyclostationary spectral analysis, *IEEE Trans. Signal Processing*, **44(11)**, 2718-2725, 1996.

[100] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, New York: McGraw-Hill, 1965.

[101] H. C. Park, D. G. Koo, J. H. Youn and J. Lee, Relay feedback approaches for the identification of Hammerstein-type nonlinear processes, *Ind. Eng. Chem. Res.*, **43(3)**, 735-740, 2004.

[102] W. Pawlak, On the series expansion approach to the identification of Hammerstein systems, *IEEE Trans. Automat. Contr.*, **36(6)**, 763-767, 1991.

[103] U. Petersohn, H. Unger and N. J. Fliege, Exact deterministic and stochastic analysis of multirate systems with application to fractional sampling rate alteration, *Proc. IEEE ISCAS*, 177-180, London, England, 1994.

[104] B. Porat, *A Course in Digital Signal Processing*, New York: Wiley, 1997.

[105] S. Prakriya and D. Hatzinakos, Blind identification of LTI-ZMNL-LTI nonlinear channle models, *IEEE Trans. Signal Processing*, **43(12)**, 3007-3013, 1995.

[106] M. B. Priestly, *Non-linear and Non-stationary Time Series Analysis*, London: Academic Press, 1989.

[107] S. J. Qin and T. A. Badgwell, A survey of industrial model predictive control technology, *Control Engineering Prac.*, **11(1)**, 733-764, 2003.

[108] A. W. Rihaczek, Signal energy distribution in time and frequency, *IEEE Trans. Inform. Theory*, **14(3)**, 369-374, 1968.

[109] D. E. Rivera, Monitoring tools for PRBS testing in closed-loop system identification, *1992 Annual AIChE Meeting*, Paper 131d, Miami Beach, FL, USA, 1992.

[110] R. S. Roberts, W. A. Brown, and H. H. Loomis Jr., Computationally efficient algorithms for cyclic spectral analysis, *IEEE Signal Processing Mag.*, **8(2)**, 38-49, 1991.

[111] B. M. Sadler and A. V. Dandawaté, Nonparametric estimation of the cyclic cross spectrum, *IEEE Trans. Inform. Theory*, **44(1)**, 351-358, 1998.

[112] H. Sakai and S. Ohno, Theory of cyclostationary processes and its application, in *Statistical Methods in Control and Signal Processing*, New York: Marcel Dekker, 1997.

[113] V. P. Sathe and P. P. Vaidyanathan, Effects of multirate systems on the statistical properties of random signals, *IEEE Trans. Signal Processing*, **41(1)**, 131-146, 1993.

[114] L. L. Scharf and B. Friedlander, Toeplitz and hankel kernels for estimating time-varying spectra of discrete-time random processes, *IEEE Trans. Signal Processing*, **49(1)**, 179-189, 2001.

[115] S. V. Schell, Asymptotic moments of estimated cyclic correlation matrices, *IEEE Trans. Signal Processing*, **43(1)**, 173-180, 1995.

[116] E. Serpedin, F. Panduru, I. Sari and G. B. Giannakis, Bibliography on cyclostationarity, *Signal Processing*, **85(12)**, 2233-2303, 2005.

[117] W. A. Sethares and T. W. Staley, Periodicity transforms, *IEEE Trans. Signal Processing*, **47(11)**, 2953-2964, 1999.

[118] R. G. Shenoy, Multirate specifications via alias-component matrices, *IEEE Trans. Circuits and Syst. II*, **45(3)**, 314-320, 1998.

[119] R. Shiavi, *Introduction to Applied Statistical Signal Analysis*, 2nd ed., San Diego: Academic Press, 1999.

[120] D.M. Sima and S. Van Huffel, Appropriate cross-validation for regularized errors-in-variables linear models, in *Proc. of the COMPSTAT 2004*, Prague, August 2004.

[121] H. C. So and Y. T. Chan, Analysis of an LMS algorithm for unbiased impulse response estimation, *IEEE Trans. Signal Processing*, **51(7)**, 2008-2013, 2003.

[122] T. Söderström and P. Stoica, *System Identification*, London: Prentice Hall, 1989.

[123] T. Söderström, U. Soverini and K. Mahata, Perspectives on errors-in-variables estimation for dynamic systems, *Signal Processing*, **82**, 1139-1154, 2002.

[124] X. Song, M. Ahmadian and S. C. Southward, Modeling Magnetorheological dampers with application of nonparametric approach, *J. Intell. Mater. Syst. Struct.*, **16(5)**, 421-432, 2005.

[125] S. X. Song, M. Ahmadian, S. C. Southward and L. R. Miller, An adaptive semiactive control algorithm for Magneto-rheological suspension systems, *J. Vibration and Acoustics*, **127(10)**, 493-502, 2005.

[126] B. F. Spencer, S. J. Dyke, M. K. Sain and J. D. Carlson, Phenomenological model of a magneto-rheological damper, *ASCE J. Eng. Mechanics*, **123(3)**, 230-238, 1997.

[127] R. Srinivasan, R. Rengaswamy, S. Narasimhan and R. Miller, Control loop performance assessment. 2. Hammerstein model approach for stiction diagnosis, *Ind. Eng. Chem. Res.*, **44(17)**, 6719-6728, 2005.

[128] T. G. Stockham Jr., High-speed convolution and correlation, *1966 Spring Joint Computer Conf., AFIPS Proc.*, **28**, 229-233, 1966.

[129] P. Stoica, On the convergence of an iterative algorithm used for Hammerstein system identification, *IEEE Trans. Automat. Contr.*, **26(4)**, 967-969, 1981.

[130] P. Stoica and T. Söderström, Instrumental-variable methods for identification of Hammerstein systems, *Int. J. Control*, **35(3)**, 459-476, 1982.

[131] L. Sun, W. Liu and A. Sano, Identification of a dynamical system with input nonlinearity, *IEE Proc.-Control Theory Appl.*, **146(1)**, 41-51, 1999.

[132] L. Sun, H. Ohmori and A. Sano, Frequency domain approach to closed-loop identification based on output inter-sampling scheme, *Proc. of the American Control Conf.*, **3**, 1802-1806, Chicago, IL, USA June 2000.

[133] L. Sun, H. Ohmori and A. Sano, Output intersampling approach to direct closed-loop identification, *IEEE Trans. Automat. Control*, **46(12)**, 1936-1941, 2001.

[134] S.W. Sung, System identification method for Hammerstein processes, *Ind. Eng. Chem. Res.*, **41(17)**, 4295-4302, 2002.

[135] C. W. Therrien, Overview of statistical signal processing, in *Digital Signal Processing Handbook*, Boca Raton: CRC Press LLC, 1999.

[136] C. W. Therrien, Issues in multirate statistical signal processing, *Proc. of the 35th Asilomar Conf. on Signals, Systems and Computers*, **1**, 573-576, Pacific Grove, CA, USA, November 2001.

[137] C. W. Therrien, Some considerations for statistical characterization of nonstationary random processes, *Proc. of the 36th Asilomar Conf. on Signals, Systems and Computers*, **2**, 1554-1558, Pacific Grove, CA, USA, November 2002.

[138] C. W. Therrien and R. Cristi, Two-dimensional spectral representation of periodic, cyclostationary, and more general random processes, *Proc. of the IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing*, **4**, 3561-3563, Orlando, USA, May 2002.

[139] N. F. Thornhill and T. Hägglund, Detection and diagnosis of oscillation in control loops, *Control Engineering Prac.*, **5(10)**, 1343-1354, 1997.

[140] N. F. Thornhill, B. Huang and H. Zhang, Detection of multiple oscillations in control loops, *J. Process Control*, **12(1)**, 91-100, 2003.

[141] C. J. Tian, A limiting property of sample autocovariances of periodically correlated processes with application to period determination, *J. Times Ser. Anal.*, **9(4)**, 411-417, 1988.

[142] L. Tong, G. Xu and T. Kailath, A new approach to blind identification and equalization of multipath channels, *25th Asilomar Conf. Signals, Systems and Computers*, **2**, 856-860, Pacific Grove, CA, November 1991.

[143] L. Tong, G. Xu and T. Kailath, Blind identification and equalization based on second-order statistics: a time domain approach, *IEEE Trans. Inform. Theory*, **40(2)**, 340-349, 1994.

[144] L. Tong, G. Xu B. Hassibi and T. Kailath, Blind identification and equalization based on second-order statistics: a frequency-domain approach, *IEEE Trans. Inform. Theory*, **41(1)**, 329-334, 1995.

[145] C. Tontiruttananon and J. K. Tugnait, Identification of closed-loop linear systems via cyclic spectral analysis given noisy input-output time-domain data, *IEEE Trans. Automat. Control*, **46(2)**, 258-275, 2001.

[146] J. K. Tugnait, Stochastic system identification with noisy input using cumulant statistics, *IEEE Trans. Automat. Control*, **37(4)**, 476-485, 1992.

[147] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Englewood Cliffs, NJ: Prentice-Hall, 1993.

[148] S. Van Huffel and Ph. Lemmerling (Eds.), *Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*, Dordrecht, the Netherlands: Kluwer Academic, 2002.

[149] A. J. van den Veen, S. Talwar and A. Paulraj, A subspace approach to blind space-time signal processing for wireless communication systems, *IEEE Trans. Signal Processing*, **45(1)**, 173-190, 1997.

[150] M. Verhaegen and D. Westwick, Identification of MIMO Hammerstein systems in the context of subspace model identification methods, *Int. J. Control*, **63(2)**, 331-349, 1996.

[151] J. Wang, T. Chen and B. Huang, Closed-loop identification via output fast sampling, *J. Process Control*, **14(5)**, 555-570, 2004.

[152] J. Wang, T. Chen and B. Huang, On spectral theory of cyclostationary signals in multirate systems, *IEEE Trans. Signal Processing*, **53(7)**, 2421-2431, 2004.

[153] J. Wang, T. Chen and B. Huang. Variability method for cyclo-period estimation of cyclo-stationary signals, *16th IFAC World Congress*, Prague, the Czech Republich, July 2005.

[154] J. Wang, T. Chen and B. Huang, Open- and closed-loop correlation analysis, *Proc. of the SICE 2005 Annual Conf.*, 230-235, Okayama, Japan, August 2005.

[155] J. Wang, T. Chen and B. Huang, Cyclo-statistic estimators for discrete-time cyclo-stationary signals, *Proc. of the SICE 2005 Annual Conf.*, 2168-2173, Okayama, Japan, August 2005.

[156] J. Wang, T. Chen and B. Huang, Cyclo-period estimation for discrete-time cyclo-stationary signals, *IEEE Trans. Signal Processing*, **54(1)**, 83-94, 2006.

[157] J. Wang, A. Sano, T. Chen and B. Huang, A new approach to blind Hammerstein identification for MR damper modeling, *Control Engineering Prac.*, submitted for publication, September 2006.

[158] J. Wang, T. Chen and B. Huang, FIR modeling for errors-in-variables/closed-loop systems by exploiting cyclo-stationarity, *Int. J. of Adaptive Control & Signal Processing*, accepted, November 2006.

[159] J. Wang, A. Sano, T. Chen and B. Huang, Blind Hammerstein identification for MR damper modeling, *2007 American Control Conf.*, accepted, December 2006.

[160] J. Wang, A. Sano, D. Shook, T. Chen and B. Huang, A blind approach to closed-loop identification of Hammerstein systems, *Int. J. Control*, **80(2)**, 302-313, 2007.

[161] P. E. Wellstead, Reference signals for closed-loop identification, *Int. J. Control*, **26(6)**, 945-962, 1977.

[162] P. E. Wellstead, Non-parametric methods of system identification, *Automatica*, **17(1)**, 55-69, 1981.

[163] G. Yang, *Large-scale magnetorheologoical fluid damper for validation mitigation: modeling, testing and control*, Indiana: The University of Notre Dame, 2001.

[164] W. X. Zheng, Least-squares identification of FIR systems subject to noise, *Proc. of the 7th Int. Conf. on Signal Processing*, **1**, 33-36, Beijing, China, 2004.

[165] Y. C. Zhu, Hammerstein model identification for control using ASYM, *Int. J. Control*, **73(18)**, 1692-1702, 2000.

[166] G. D. Zivanovic and W. A. Gardner, Degree of cyclostationarity and their application to signal detection and estimation, *Signal Processing*, **22(3)**, 287-297, 1991.